DOCUMENT RESUME

ED 060 681 FL 002 814

AUTHOR	Levin, Lennart; Olsson, Margareta
TITLE	Learning Grammar: An Experiment in Applied Psycholinguistics Assessing Three Different Methods of Teaching Grammatical Structures in English as a Foreign Language.
INSTITUTION	Gothenburg School of Education (Sweden).; Gothenburg Univ. (Sweden). Dept. of English.
SPONS AGENCY	National Swedish Board of Education, Stockholm.
REPORT NO	RBull-7
PUB DATE	Jan 71
NOTE	184p.
EDRS PRICE	MF-$0.65 HC-$6.58
DESCRIPTORS	Applied Linguistics; Attitudes; Cognitive Processes; Computers; Contrastive Linguistics; *Educational Experiments; Educational Strategies; *English (Second Language); *Grammar; Instructional Materials; Language Patterns; Language Research; Lesson Plans; Program Evaluation; Psycholinguistics; *Second Language Learning; Statistical Analysis; Structural Analysis; *Teaching Methods; Testing

ABSTRACT

Project 5 of the GUME study on foreign language teaching methods continues the work of the previous projects concerning the relative effectiveness of a habit-formation (implicit) method and a cognitive-code learning method, with explanations in either the source or target language. The educational problem of Project 5 is teaching the English passive voice to Swedish students in Class 8. The teaching procedures and design differ slightly from the first three projects. The time for explanations varies between the source-language and target-language methods. The source-language method uses comparisons with corresponding structures in the source language where applicable. No significant differences in the learning effects of the three methods are evidenced; interesting trends, however, are noted and discussed. This report presents details on project design, lessons, and evaluation instruments. A statistical description of the experimental population is provided along with a statistical analysis of the results and a correlation study of the variable factors. A discussion of the results and their implications follows. A bibliography is included and appendixes provide details on the English grammatical explanations and on some of the evaluation instruments used. Related documents can be found through the following reference numbers: Project 1, ED 034 172; Project 2, FL 002 818; Project 3, FL 002 819; Project 4, ED 045 969; Adult Project, FL 002 868; statistical synopsis of Projects 1-3, FL 002 816. (VM)

ED 060681

The GUME Project

L E A R N I N G G R A M M A R

An Experiment in Applied Psycholinguistics

Assessing Three Different Methods of Teaching Grammatical Structures in English as a Foreign Language

by

Lennart Levin and Margareta Olsson

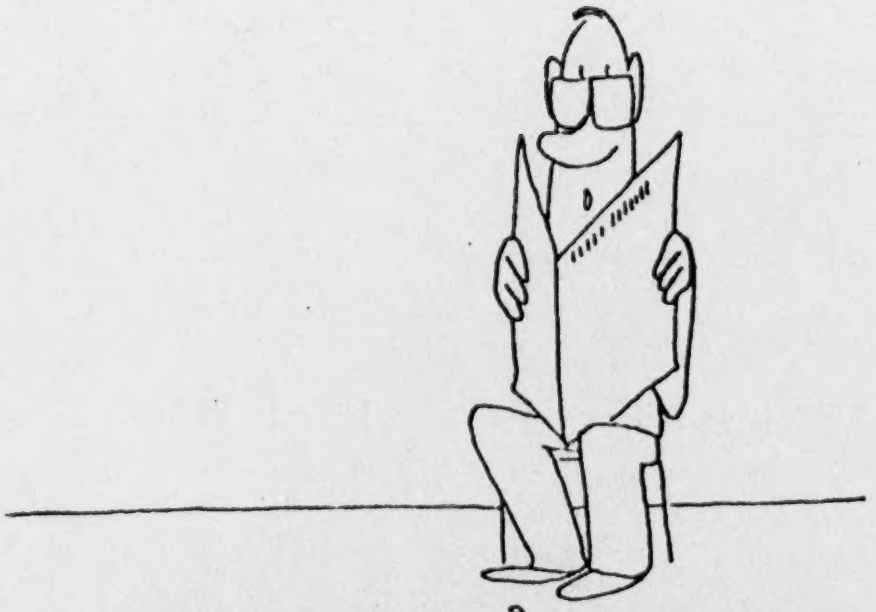
U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Gothenburg School of Education
Department of Educational Research
and
Department of English

University of Gothenburg
Department of English

January, 1971

FL 002 814



DET FINNS DAGAR DÅ TIDNINGARNA INTE INNEHÅLLER EN ENDA NY REVOLUTIONERANDE UNDERVISNINGS-METOD.

There are days when you can´t find a single new revolutionary teaching method in the newspapers.

Within the GUME project (Göteborg, UndervisningsMetod i Engelska = Gothenburg/Teaching/Methods/English) earlier studies showed no significant differences in learning effects between different methods of teaching English.

Together with the GUME 4 study, published in December 1970 in this report series, it is a direct continuation of the earlier studies. Modifications in design, teaching strategies, etc., have been made in order to increase the probability of detecting true differences, if such exist, between methods. As in the previous experiment, the three methods being compared are: the Implicit method, the Explicit-English method, and the Explicit-Swedish method. In all the methods the students have systematized drills; in Ee and Es the students have analysis and explanations as well. In Ee these explanations are given in the target language and in Es in the source language. In Es comparisons are also made with the corresponding grammatical structures in Swedish.

The experiment took place in grade 8 of the comprehensive school. The specific grammatical structure taught is the passive voice. The experimental population consists of 12 school classes belonging to the advanced course and 12 classes representing the easier course in English. Within each course the classes were randomly assigned to teaching method.

Main effects are investigated by analysis of covariance and interaction effects by analysis of variance (two-way classification). Individual scores and, in some analyses, school class means are used as units of analysis. Various measures of progress during the experiment are used in the comparisons.

ACKNOWLEDGEMENTS

This study forms the fifth part of the GUME-project. The central concern of the project is to investigate learning effects when different teaching techniques are used in the foreign language instruction.

This is an interdepartmental undertaking and the project members have enjoyed the stimulating influence not only of Professor Alvar Ellegård, head of the English Department of the University of Gothenburg, but also of Professor Karl-Gustaf Stukát, head of the Department of Educational Research of the Gothenburg School of Education, which has been a great asset to the project as a whole.

The present study was supported by grants from the National Board of Education, bureau L 4, and we are greatly indebted to this institution. We are also grateful to the headmasters who gave us access to their schools and to the many classroom teachers who with goodwill and patience endured the extra work and instructional restraints imposed upon them by participating in the project.

Our thanks are also due to Per Högberg, our programming expert, and to Kerstin Davidsson, who converted our anything but successful attempts at type-writing into well-arranged and tidy pages.

N O T E

The present investigation, GUME 5, is together with the study reported on in the previous issue of this report series, a direct continuation of earlier studies on foreign language teaching methodology.

It was a delicate task to decide whether the research represented by the two part projects should be reported in one or in two separate volumes. Much was speaking for a single report, above all the similarity in design, statistical treatment and general hypothesis to be tested. However, much was also speaking against it: different grammatical structures were taught in the two investigations, the experimental populations were of different ages, there were variations in the number of lessons during the experimental instruction. After careful consideration of the pros and cons of the two alternatives, we decided to publish two reports.

In order to make the reader of the present volume relatively independent of the GUME 4 report, a few sections or chapters contained in the latter have also been included here. They are:

INTRODUCTORY NOTE ON THE TREATMENT OF STATISTICS
BACKGROUND (*in part*)
THE PENNSYLVANIA PROJECT CONTINUED
EXPERIMENTATIONS IN A FIELD SETTING - SOME REFLECTIONS (*with minor alterations*)

TABLE OF CONTENTS

ACKNOWLEDGEMENTS i

NOTE ii

TABLE OF CONTENTS iii

List of Figures vii

List of Tables viii

INTRODUCTORY NOTE ON THE TREATMENT OF STATISTICS xi

BACKGROUND 1

Earlier GUME Activities 1

Total GUME Activities 4

THE PENNSYLVANIA PROJECT CONTINUED 6

EXPERIMENTATION IN A FIELD SETTING - SOME REFLEXIONS 16

Comparative Experiments - Pros and Cons 16

The GUME Project - Some Comments 18

Hypothetical Treatment Effects 19

METHODS IN FOREIGN LANGUAGE TEACHING 22

Current Methods and Their Historical Background 22

The Authorized Curricula for Schools 24

Towards a Definition of the GUME Methods 25

The Implicit method 26

The Explicit methods 26

Stages in Foreign Language Learning 27

Debatable points 29

Recent Methodological Discussion 30

A DESCRIPTION OF THE LESSON SERIES 34

Lesson Design 34

Written activities 35

Oral drills 35

The Explanations 37

Instructional Content per Lesson 38

Survey of the Lesson Content 39

THE GUME 5 PROJECT A DESCRIPTION OF THE LAY-OUT 41
Objectives 41
The Present Investigation in Relation to Earlier Results . 41
Time Schedule 43
The Teaching Methods 44
The Experimental Population 44
Assignment to Treatments 45
Drop-out Rate 45

EVALUATION INSTRUMENTS 47
The Criterion Test 47
Description of the Sub-tests 48
Part 1 48
Part 2 48
Part 3 48
Part 4 48
Part 5 48
Part 6 49
Reliability 49
Validity 50
The Pupil Attitude Test 51
The Teacher Attitude Test 52
The Standardized Test in English 52

PACT 53
General Scholastic Aptitude (DBA) 56
Other Measures 56
Social class 56
Grades 56
Grades German and French 57
The Reliability of Subjectively Scored Sub-tests 57
Comments on an Oral Test 57

DATA TREATMENT 59
The Statistical Program 59
Experimental Design 60
Investigation of Main Effects 60
Investigation of Interaction Effects 61
Retention 61
Various Measures of Progress 62

STATISTICAL DESCRIPTION OF THE EXPERIMENTAL POPULATION 63
Attendance 63
Boys/Girls Ratios 63
Assignment to Treatments 64
Social Class 65
Scholastic Aptitude 68
Grades 69
The Standardized Test in English 70
Sex Differences in Background Variables 71
Characteristics of the Treatment Groups 73

MAIN RESULTS 74
Overall Progress during the Experiment 74
Progress - Main Effects 76
School class progress 76
Individual progress per method 78
Progress - Interaction 81
School Class Data Analysed 83
Additional Measures of Progress 84
Results on the Different Part Tests 88
Drop-outs 91
Progress from a Different Point of View 93

CORRELATION STUDIES 94
Intercorrelations between the Main Variables (the sk sample) 94
Progress 94
Pupil attitudes 94
Social class 94
PACT 94
The DBA variables 96
The grades 96
Grades German and French 96
Intercorrelations between the Main Variables (the ak sample) 96
Progress 96
Pupil attitudes 98
Social class 98
PACT 98
The DBA variables 98

The grades 98
Grades German 98
Pre-test and Post-test Correlations (the sk sample) 99
Pre-test and Post-test Correlations (the ak sample) 99
Pre-, Post-test and Progress Correlations 102
THE PUPIL ATTITUDE TEST 103
Attitudes towards the explanations 106
General interest in English 106
THE TEACHER ATTITUDE TEST 107
DISCUSSION OF RESULTS 117
SUMMARY 120
BIBLIOGRAPHY 124
APPENDICES A-F

List of Figures.

Fig. 1: A Survey of GUME Research Activities 1968-1971 5

Fig. 2: Chronology of GUME 5 43

Fig. 3: Distribution of Pupils (%) according to Social Class and Course in English 67

Fig. 4: Distribution of Progress Scores for sk and ak 75

Fig. 5: The Progress of the 24 Experimental Classes (12 sk and 12 ak Classes) 77

List of Tables

Table 1: Allocation of Time within the three Strategies 34

Table 2: Allocation of Time per Lesson within the Three Strategies 36

Table 3: Occurence and Length of the Explanations in the Ee and Es Strategies 38

Table 4: Number of Pupils Participating in the Study 46

Table 5: Attendance of the Experimental Population (sk + ak) during the Series of Lessons 63

Table 6: Number of Boys and Girls in the Experimental Group (sk + ak) and in the Total Population 64

Table 7: Distribution of Pupils according to Teaching Methods (sk) 65

Table 8: Distribution of Pupils on Teaching Methods (ak) 65

Table 9: Distribution according to Social Class (absolute numbers) sk + ak 66

Table 10: Distribution according to Social Class (percentages) sk + ak 66

Table 11: Means and Standard Deviations on the DBA-test. (The Parts in Stanine Points, the Total in T-scores) 68

Table 12: Means and Standard Deviations for English, Swedish, and Mathematics 69

Table 13: Means and Standard Deviations on the Standardized Achievement Test in English 70

Table 14: Means and Standard Deviations for Boys and Girls in Certain Variables (sk) 71

Table 15: Means and Standard Deviations for Boys and Girls in Certain Variables (ak) 72

Table 16: Analyses of Variance (One-Way) of Treatment Group Means in Certain Variables; sk 73

Table 17: Analyses of Variance (One-Way) of Treatment Group Means in Certain Variables; ak 73

Table 18: Pre-test, Post-test, and Progress Means and Standard Deviations. sk; N = 235 74

Table 19: Pre-test, Post-test, and Progress Means and Standard Deviations. ak; N = 152 76

Table 20: Analysis of Covariance, sk. Dependent Variable: PROGRESS. Covariates: DBA, PACT, the Pre-test, the Standardized Test in English, and the Weighted Sum of the Four 79

Table 21: Analysis of Covariance, ak. Dependent Variable: PROGRESS. Covariates: DBA, PACT, the Pre-test, the Standardized Test in English, and the Weighted Sum of the Four 80

Table 22: Analyses of Covariance, sk and ak. Dependent Variable: POST-TEST. Covariate: Pre-test 80

Table 23: Distribution of sk and ak Pupils on Three Levels of Scholastic Aptitude According to the DBA Test 82

Table 24: Analysis of Variance (Two-Way); sk + ak. Dependent Variable: PROGRESS 82

Table 25: School Class Means (ak) for the Pre-test and the Post-test. N = 12 83

Table 26: Analysis of Covariance of School Class Means (N = 12) Dependent Variable: the POST-TEST. Covariate: the Pre-test 84

Table 27: School Class Means for: Actual Improvement x 100/ / Possible Improvement; sk, N = 12 84

Table 28: Analysis of Variance (One-Way) of School Class Means on the Variable: Actual Improvement x 100 / / Possible Improvement; sk, N = 12 85

Table 29: School Class Means for: Actual Improvement x 100 / / Possible Improvement, ak, N = 12 85

Table 30: Analysis of Variance (One-Way) of School Class Means on the Variable: Actual Improvement x 100 / / Possible Improvement, ak, N = 12 85

Table 31: School Class Means on the Variable: Progress x 100 / / Pre-test; sk, N = 12 86

Table 32: Analysis of Variance (One-Way) of School Class Means on the Variable: Progress x 100 / Pre-test; sk, N = 12 86

Table 33: School Class Means on the Variable: Progress x 100 / / Pre-test; ak, N = 12 87

Table 34: Analysis of Variance (One-Way Classification) of School Class Means on the Variable: Progress x 100 / / Pre-test; ak, N = 12 87

Table 35: Results on the Part Tests per Method; sk 88

Table 36: Results on the Part Tests per Method; ak 90

Table 37: Means and Standard Deviations for the Experimental Population and the Drop-outs (sk) 91

Table 38: Means and Standard Deviations for the Experimental Population and the Drop-outs (ak) 92

Table 39: Intercorrelations between the Main Variables; sk, N = 235 95

Table 40: Intercorrelations between the Main Variables; ak, N = 152 97

Table 41: Pre-test and Post-test Correlations; sk N = 235 ... 100

Table 42: Pre-test and Post-test Correlations; ak N = 152 ... 101

Table 43: Pupils' responses on the Questionnaire (Questions with fixed response alternatives). sk and ak 103

Table 44: School Class Means for Attitudes to the Experiment; sk N = 12 104

Table 45: Analysis of Variance of School Class Means for Attitudes to the Experiment 104

Table 46: Distribution of Responses to Item No. 12 of the Questionnaire Concerning the Explanations 106

INTRODUCTORY NOTE ON THE TREATMENT OF STATISTICS

The study dealt with in the present report is an interdepartmental (tvärvetenskaplig) undertaking, one of the authors representing English as an academic discipline and a school subject, one representing pedagogy as an academic discipline and educational research and statistics as theoretical background. We have written the report with two quite distinct groups of readers in mind: teachers of English and educational researchers. The former group normally has little training in statistics and has a tendency to shy away from figures, the latter has training in this field and is perhaps more used to reading reports like the present one. This has caused problems in writing the report.

What we have tried to do is the following. We have used ordinary statistical methods and give as much information and as many tables as will hopefully satisfy the second group of our intended readers. But we have also tried to arrange the tables so as to facilitate the reading of them for the first group of readers. The language teacher with little training in statistics is recommended to study columns and tables of means and standard deviations, and N (see below). In commenting on our tables we have not always limited ourselves to conclusions and discussions of these but have also tried to explain how we arrived at these conclusions, how the figures ought to be understood, what size a certain figure must reach to be "significant", etc. We hope that those readers who find these comments superfluous will understand the pedagogical *raison d'être* for them and will just skip them.

For the convenience of the reader with little statistical training some frequent symbols and terms are explained below. In almost every case the explanation is an attempt at giving the general idea or practical use of a symbol rather than an adequate or in all respects logical definition of it.

N — The *number* of pupils for which a certain measure is given.

$\bar{x}$ — The *arithmetic mean* of a group.

s — The *standard deviation*, i.e. a measure of the extent to which the scores for a certain group vary. The larger the s, the more heterogeneous the group. A single s does not carry much meaning; the measure should be used for comparison with other s's.

t — This value indicates whether a difference between the means of two groups is "statistically significant" or whether it can be explained as a chance occurrence. As far as the analyses in the present report are concerned the critical t-value is 1.96, i.e. when *t* is equal to or greater than 1.96, the difference under investigation is considered a real, non-chance difference.

F — F, or the F-ratio, is used for the same purposes as *t*. However, F is the relevant characteristic when more than two means are compared. Since three teaching methods are being compared in the present study, F appears quite often in our tables. The corresponding critical value for interpreting differences as true differences lies around 3.00; this figure varies a little depending on the number of pupils.

T-scale — A scale with a theoretical mean of 50 and a standard deviation of 10. The scores on a certain test, whatever its $\bar{x}$ and s, can be transformed into T-scores.

Stanine scale — A 9-point scale with a theoretical mean of 5 and a standard deviation of 2. In contrast to the T-scale, the stanine scale has a so-called standardized (normalized) distribution of scores. Scores on a test may be transformed to stanines by giving the top and bottom 4 % of the pupils 9 and 1 points respectively, the next 7 % at each end 8 and 2 respectively, thus: 9 (4 %), 8 (7 %), 7 (12 %), 6 (17 %), 5 (20 %), 4 (17 %), 3 (12 %), 2 (7 %), 1 (4 %).

Analysis of variance — The method is used for comparing the means of three or more groups which have been exposed to different treatments. Do the groups respond in different ways, i.e. are their means statistically different? In this sort of analysis, the variation in scores between groups and within groups are considered in relation to each other. For true differences between group means to exist, it is necessary for the variation in scores *between* groups to be greater than the variation *within* groups. This sort of analysis yields an F-ratio (see above).

Analysis of covariance — The same as the above method with the addition that the groups' standing on essential background variables is taken into account. For instance, if three groups are to be compared with respect to learning effects and the groups differ substantially in intelligence, it is very probable that the group having the brightest children (and not necessarily the children exposed to the "best" method) would come out as the best. In an analysis of covariance, differences of this sort are equalled out statistically. This analysis also yields an F-ratio.

Adjusted means Refers to analyses of covariance. The means of the groups being compared are adjusted for variation between the groups in background variables. Briefly, if three groups were to rank A > B > C in a teaching experiment and their values in the background variable, say intelligence, also ranked A > B > C, the adjusted means would be equal for the three groups. Thus, when original differences between the three groups were taken into consideration, differences obtained after the teaching experiment disappeared.

X^2 (Chi^2) A value used to indicate whether the answers on, for instance, a questionnaire are evenly distributed among the response alternatives. It is used to investigate if the particular distribution of answers (given by a group of individuals) is in accordance with an expected distribution and if a deviation in this respect is so small that it might be explained as a chance occurrence. The differences between observed and (theoretically) expected frequencies add up to a so-called X^2 -value; the higher this value, the more probable is the conclusion that the group (of pupils, etc) under consideration deviates significantly from "the norm".

BACKGROUND

Earlier GUME Activities

The present report describes further research on the teaching of English as a foreign language by members of the so-called GUME project. The work should be viewed against the background of four separate reports, published in 1969 (see special section of the bibliography, page 124) and describing teaching method comparisons performed so far. For readers not familiar with the publications just mentioned, a brief resumé may be in order:

Three parallel studies, identical in design, were carried out in order to investigate three different methods of teaching grammatical structures in English as a foreign language. The studies were performed during the autumn term of 1968 and the spring term of 1969. Three different areas of English grammar that are known to cause Swedish students difficulty were selected for investigation:

GUME 1 The do-construction
GUME 2 The some-any dichotomy
GUME 3 The passive voice

The three methods of instruction (independent variables) investigated in each of the experiments were:

Im *The Implicit method*, where the students had systematized drills but no analysis or explanations of the grammatical structures involved.

Ee *The Explicit-English method*, where the students had systematized drills and, in addition, analysis and explanations in the target language (English). The time allotted to the explanations was taken from the drills.

Es *The Explicit-Swedish method*, where the students had systematized drills and, in addition, analysis and explanations in the source language (Swedish); comparisons with corresponding structures in Swedish were also made. The time allotted to the explanations was taken from the drills.

In each part project 18 school classes took part, 6 per teaching strategy. Of these 6 classes, 4 represented the advanced course

(särskild kurs, abbreviated sk) and 2 the easier course (allmän kurs, abbreviated ak). Thus the total GUME project contained 54 classes, of which 36 were in sk and 18 in ak. The school classes, representing a wide geographical variation within the Gothenburg area, were randomly assigned to the teaching methods.

For each part project 3 lesson series (Im/Ee/Es) were constructed, each consisting of 6 lessons. In order to control the teacher factor "canned" lessons were used throughout the experiment. The students listened to the programs via headsets with induction receivers. Magnetic wires were installed and tape-recorders used in every classroom; this simple arrangement comes close to a language lab as far as sound quality is concerned.

Within each part project, the pupils' progress was measured by a criterion test, designed to correspond to the specific objectives of the part project in question. That is to say, the same test was administered as Pre-test before and as Post-test after the experiment, the difference between the two being the Progress score for each pupil. The identical test was also administered as Re-test approximately one month after the experiment in order to measure retention.

The pupils' attitudes to various aspects of the study were collected by means of a questionnaire.

Since the treatment groups within each experiment were not experimentally controlled, statistical control was undertaken by means of analysis of covariance. The covariates resorted to were "general intelligence" (the verbal, inductive and spatial factors of an IQ test frequently used in Swedish schools), grades in English, Swedish and Mathematics, and in some analyses Pre-test scores. Partly the analyses were made with Progress scores as the dependent variable and partly with Post-test scores as the dependent variable.

In the various statistical analyses the experimental population was divided according to two principles: in one type of analysis sk and ak were treated separately, in another the population was divided into three equal parts according to IQ scores, the Upper, Middle and Lower third. In the latter case analyses of variance (two-way classification) were performed in order to investigate interaction between ability level and teaching method.

More detailed information about the statistical treatment of GUME 1-3 will not be given in this connection, suffice it to say that a total of 60 (sixty) analyses of covariance and variance were performed.

In GUME 1-3 two statistically significant differences were obtained, which is less than could be accounted for by mere chance even if the null hypothesis (no difference between treatments) were true. Nor was there any evidence of interaction between ability level and teaching strategy in the study.

Thus the GUME 1-3 experiments have not shown that any differences are produced by the three teaching methods.

The GUME 4 project, of which the field work phase occurred simultaneously with that of the present project (see figure 1, p. 5 below), has been reported on in December 1970 (see special section in the bibliography, p. 124). In comparison with earlier GUME studies, more grammatical structures were taught in the GUME 4 project, the duration of the experiment was doubled and new, hopefully more optimal, explanations, were formulated for the Explicit groups. However, in all essentials the main results were identical with those obtained earlier: no differences in learning effects were produced by the three teaching methods, nor was there any evidence of interaction between the pupils' ability level (as measured by a test of scholastic aptitude) and teaching method. Against the background of these results it was questioned whether the intense debate on foreign language teaching methodology in Sweden during 1969 to 1970 has not been on the wrong track; new areas of investigation, probably more rewarding than those of teaching method comparisons, should be discussed.

It is sometimes argued that "insignificant" results like those obtained in GUME 1-3 have low social utility (Anderson, 1969) since they do not provide much support for people involved in the production of teaching materials.

In the three studies referred to, however, the main concern was the basic problem of whether explanations facilitate learning rather than the production of materials. Consequently the lessons were designed to provide an answer to the basic research question without necessarily coming close to "ordinary" lessons. Even so, no differences were found between the three teaching methods compared. (If significant

differences had appeared, they would still have been of limited interest *with respect to the production of materials.*)

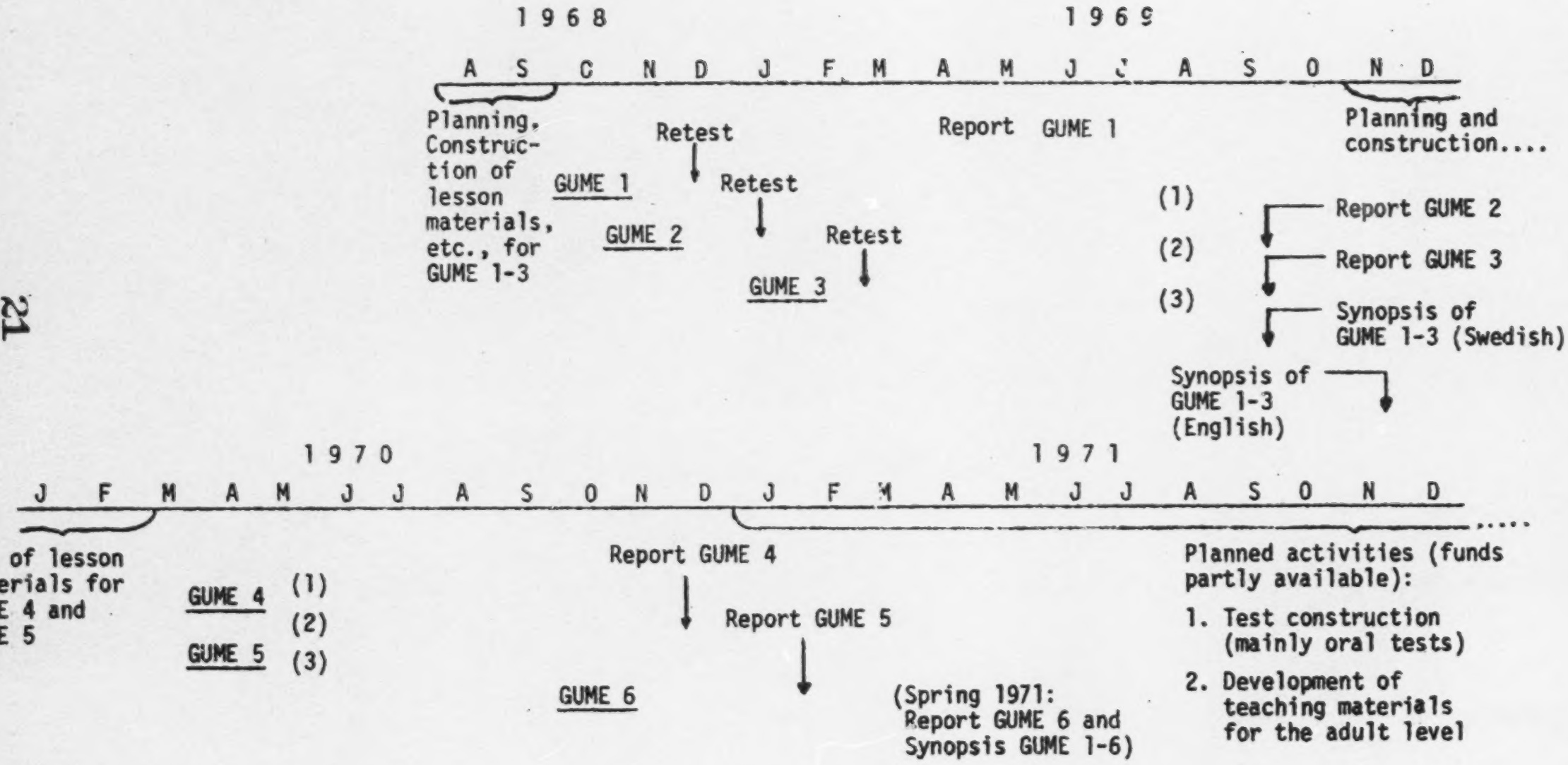
Findings like those just reported are not uncommon in educational research (Stephens, 1967). True differences between methods may have escaped detection because the experiments lacked statistical power (Stanley, 1970) or because of deficiencies in the planning and execution of the studies. There is also the possibility that no true differences between the methods exist, though this can never be proved.

Total GUME Activities.

So far the reader has become acquainted with the four first part projects, GUME 1-4. Together with the present investigation they cover grades 6,7, and 8 of the Swedish comprehensive school in what was considered an urgent research undertaking. However, under the assumption that foreign language teaching, and learning, may function differently at the adult level, still another part project was performed in 1970. This project, GUME 6, will be reported on during the spring of 1971. The following brief mention of it is indented to complete the picture of the total GUME activities.

GUME 6 is performed at the adult level. The strategy adopted is to compare two methods only, one of an audiolingual kind with numerous structure drills and no explanations or generalizations, and one with very few drills but with explanations in the source language. The two methods are intentionally made more distinctly different than, for instance, Im vs. Ee/Es in the earlier GUME experiments. Fig. 1 gives a survey of the GUME studies, performed in the past as well as planned for the future. At one point a clarification is necessary; the figures 1, 2, and 3, appearing in two positions, indicate that the criterion tests used in GUME 1, 2, and 3 respectively were administered in control classes at the beginning and at the end of the school year. The purpose was to find out to what extent the structures taught during the GUME experiments are actually learnt in a school year without the teachers' paying special attention to those structures. Progress in the control classes will be commented on in the present report (p. 93).

Figure 1: A Survey of GUME Research Activities 1968-1971.


1968
1969
A S C N D J F M A M J J A S O N D
Planning, Construc-tion of lesson materials, etc., for GUME 1-3
GUME 1
Retest
GUME 2
Retest
GUME 3
Retest
Report GUME 1
Planning and construction....
(1)
(2)
(3)
Report GUME 2
Report GUME 3
Synopsis of GUME 1-3 (Swedish)
Synopsis of GUME 1-3 (English)
1970
1971
J F M A M J J A S O N D J F M A M J J A S O N D .....
... of lesson materials for GUME 4 and GUME 5
GUME 4
GUME 5
(1)
(2)
(3)
Report GUME 4
Report GUME 5
GUME 6
(Spring 1971: Report GUME 6 and Synopsis GUME 1-6)
Planned activities (funds partly available):
1. Test construction (mainly oral tests)
2. Development of teaching materials for the adult level

THE PENNSYLVANIA PROJECT CONTINUED

The largest undertaking in recent years in the field of educational research concerning the teaching of foreign languages is the Pennsylvania study. The GUME project is a similar enterprise although on a much more modest scale, smaller in scope and personnel. We have studied the Pennsylvania reports carefully and tried to learn both from those parts of the design and evaluation which are worthy of imitation, and from the mistakes and shortcomings. In an earlier report (Levin, 1969, p. 6 ff) we gave a commented outline of the study, including what had been reported by September, 1969. The debate in USA has been lively, and since much of the criticism levelled at the Pennsylvania Project might be directed at us, we have considered it worth-while to give a fairly extensive survey of this debate and its main arguments. This might seem to be somewhat outside the scope of the present report, but the survey has been written with the direct bearing on the GUME project of the debate in view, even if this is not explicitly pointed out more than once or twice.

When the outline of the Pennsylvania Project, given in the synopsis of the earlier GUME studies (Levin, 1969, p. 6 ff) was written, the results of the two first years' studies (as reported in Smith-Berger, 1968, and Smith-Baranyi, 1968) were available. As a matter of fact, a preliminary report on the third year follow-up was also at hand; however, we then abstained from commenting on more than levels I and II, i.e. the first two years of investigation. Since that time a supplementary report (Smith, 1969a), covering the third and fourth year results as well as complementary statistical treatment of level I and II data, has become available. Various members of the GUME project have also had the privilege of personally obtaining any information desired from Dr Philip Smith, Jr., the project coordinator.

The reader is referred to the above mentioned synopsis for an outline of the Pennsylvania project, its objectives, research design, etc. (Of course we agree with those reviewers of the Pennsylvania project who recommend interested readers to consult the full reports. Any brief critique fails to do justice to the full scope of the findings). The following sketch is for the benefit of readers not acquainted with the Pennsylvania project.

The main purposes were to investigate which of three foreign language strategies was most effective and to determine which of three language laboratory systems was best suited, economically and instructionally, to the development of pronunciation and structural accuracy. The three teaching methods compared were the Traditional Method (TLM), the Functional Skills Method (FSM), and the Functional Skills + Grammar Method (FSG); the three laboratory systems compared were Tape Recorder only (TR), the Audio-Active system (AA), and the Audio-Active-Record system (AAR). The intact school class was the experimental unit. Class assignment was random only across the two functional skills methods (in the case of TLM only teachers who had expressed a preference for that method were assigned to it). The original (= first year's) population consisted of 104 school classes (61 French, 43 German) from nearly as many schools, representing a great geographical variation within the state of Pennsylvania. Of the original 104 classes, 61 remained throughout the second year. After two years, the main finding, obviously not expected by the profession, was that no significant differences existed among strategies on all skills except reading (TLM $>$) as measured by contemporary tests. Nor did the language laboratory of any type, used twice weekly, have any discernible effect on achievement. The criticism that we ventured to pass in our previous report on the research performed thus far (levels I and II) may be summarized thus:

1. The non-random assignment of classes to treatments (in the case of TLM) is a potential source of error in that teacher preference may reflect belief in that strategy, which will breed more enthusiasm for the work and hence increase the chances of better results.

2. The two "Functional Skills" methods do not seem to be very distinct; considering the diffuse difference between FSM and FSG one might suspect that the experiment is, in reality, a comparison between one traditional and *one* audio-lingual method.

3. No special course material was constructed. The project staff chose five French and four German textbooks out of the twenty-seven which are commonly used and decided which were to be used in each method. Most teachers were thus left with a limited choice. No maximum pensum to be read was established; the different classes could (and did!) cover different amounts of text. Thus

text materials chosen as well as rate of progress in the text-books are possible sources of variation. (As a matter of fact, during the first year, TLM classes covered almost three times as much text as did the FS classes.)

4. An outdated version of the MLA Cooperative Tests (1939-41), apparently favouring TLM classes, was used in one phase of the study.

(A Swedish reader should be aware that the experimental population, compared to Swedish circumstances, was a very select group since only 17-20 % take a foreign language in Pennsylvania; thus even the "low IQ group" would be part of the upper IQ third of the GUME population.)

In the final report (Smith, 1969a)it becomes evident (p. 23) that too few French students remained in the Traditional experimental treatment after three years for meaningful comparisons to be made with Functional Skills classes. The third year summary reads (p. 41): "A sufficient number of German students remained available to the project staff through Level III to support the conclusions drawn after Levels I and II: 'Traditional' students equaled or significantly exceeded the achievement of 'Functional Skills' students on the MLA Cooperative Classroom Listening and Reading Tests". It should be mentioned that two more conclusions were forwarded, one concerning correlations between measures of teacher proficiency and school class achievement, and one concerning student opinion measures; however, our concern here is with the main results.

Complete data extending over a full four-year period was obtained on 92 students, 72 German and 20 French, i.e. 2 % of the original population. The German students were quite evenly distributed among the three strategies: TLM: 27, FSM: 24, FSG: 21. This sample permitted the computation of an analysis of covariance using the pre-experimental Modern Language Aptitude Test as a covariate. For the French students no such investigation of main effects was possible. The fourth year summary reads as follows (p. 44): "Level IV results support earlier findings that there is no advantage favoring Functional Skills classes in performance on tests designed to measure functional skills. IQ seems to be the best predictor of long-range student foreign language achievement within the secondary school setting". The final report also contains additional information and analyses of the first and second years of study and, most interestingly, a "Condensation of Discussion Conference Proceedings".

The following section is a review of reviews; in the case of the Pennsylvania study, the results of which stirred up emotions and initiated a lot of reviews, this may be a contribution in its own right.

The reviews we shall comment on here are Carroll's (1969) and Wiley's (1969) in the December issue of *Foreign Language Annals*, 1969, and various articles in the now famous October issue of the *Modern Language Journal*, 1969.

In our own review in the previous report (Levin, 1969, p. 6) we stated that the Pennsylvania project would probably become a classic, considering the investment in people and money. Dr. Philip Smith Jr. gives the following factual information on the scope of the investigation (1969c, p. 2): ..... "four thousand two hundred students in one hundred and thirty-two classes representing an investment of three hundred and fifty thousand dollars and over a thousand pages of written materials , ..." Similarly, Carroll says (p. 214): "The Pennsylvania Foreign Language Research Project will undoubtedly go down in the annals of foreign language teaching research as one of the classics. In size, scope, carefulness of experimental design, and importance of results it is unmatched by any previous study of its kind. It has already attracted wide attention because of the apparent discrepancy between its findings and the outcomes that current thinking about foreign language teaching might have led one to expect or to hope for". As the last sentence indicates Carroll is obviously assuming that the profession at large would expect results favouring the audio-lingual methods rather than the traditional. Carroll, although professing that he does not intend to choose sides in the debate, admits his own bias towards a "cognitive code-learning" approach, which undoubtedly has more in common with the TLM than the other two methods in the Pennsylvania study. Perhaps it is this inclination that causes him to take the results, at least to some extent, at their face value (p. 214): "In brief, it (the study) seems to tell us that the 'audio-lingual' emphasis of current FL teaching philosophy is in some way misguided".

Carroll is almost laudatory with respect to the experimental design of the study. "In fact, it is one of the few large-scale studies that has well observed the canons of scientific educational research" (p. 215). This is in agreement with Wiley who states (p. 211): "(In spite of

these criticisms) the design and its implementation were excellent in comparison to other evaluation studies in that no attempt at random assignment of relevant units to treatments is usually made". The following quotation is intended to illustrate the inconsistency between different reviews by qualified researchers (Aleamoni & Spencer, 1969, p. 421): "The study appears to fall more into the category of an *ex post facto* research design while professing to be an experimental design. The *ex post facto* research design does not allow testing for treatment effects but, instead, only permits comparisons between groups, etc., on common variables. In the case of the Pennsylvania Project, data could be collected under this model to determine differences of student achievement in existing but varying classroom conditions, *but the results would not indicate what, if any, effect the classroom conditions had on student achievement*" (italics ours). If this critique were valid, and our own belief is that it is not, the results of the study would be highly suspect.

To return to Carroll, he makes the observation (p. 235) that "the 'Traditional' method used in the study was apparently, in most cases, a 'traditional-modified' method which exposed the student to a considerable amount of spoken language (cf p. 30 below). The most misleading thing about the publicity that has attended the study is the use of the word 'traditional', which will be interpreted by the casual reader as meaning a form of FL instruction that may have been prevalent forty years ago but that hardly has a place in to-day's schools". It is unfortunate that the observation scales used for describing classroom activities were constructed so as not to make control of adherence to method by teachers possible (a fact which has been pointed out by several reviewers); as Carroll observes, TLM students obviously used oral language more than they were supposed to (218). If this observation by Carroll is correct, and similarly, if our own statement concerning the diffuse differences between FSM and FSG is correct, then, which *were* the methods being compared in the Pennsylvania project? If we have stressed this point strongly here, it is because we have become aware, during the course of our own work, of the difficulty of keeping the methods distinct (though this must be far more easy in the case of "canned" materials).

Some of the criticisms that Carroll passes on the study are:

Too few classes remain in some of the strategy-system cells for statistical inferences to be made.

The text used, rather than the method, may explain some of the main effects (in Carroll's terminology, the text is a "stowaway variable").

Control of vocabulary load should have been made in the case of the criterion tests.

Sampling of classes was not strictly random.

Some selectivity in the reporting of data can be noticed. ("As this critique demonstrates, the readers of a statistical report sometimes find it necessary to refer to data that the investigators may not think worth reporting", p. 221).

No rationale was given for the choice of covariates.

No two-way analyses of variance were made in order to investigate interaction between strategy and ability.

The tests of "teacher proficiency" were in no sense intended to measure actual *ability to teach a foreign language*; apart from the misleading term, Carroll criticizes the statistical treatment of "teacher data" for being incomplete.

Our review of Carroll's review has been severely selective in that we have hardly made justice to his fundamentally positive attitude to the research completed by the Pennsylvania project staff. Our negative bias has had one aim: to provide the reader and ourselves with a "check-list" when contemplating the present report.

A final quotation from Carroll's review (p. 234): "I do believe that the findings of the study with regard to teaching strategies and laboratory systems are sufficiently solid and replicable to prompt us to rethink methods and objectives in foreign language teaching".

Wiley's review concentrates on the design and the statistical treatment of the results. The most serious defect in the design, according to Wiley, is the non-random assignment of classes to treatments. He points out that the average IQ in schools which had a language laboratory might be different from the IQ in schools without these facilities; thus *presence or absence* of a language laboratory might be associated with background variables. Because of this possibility it

is unfortunate that no analyses of Pre-test data are reported so that this suggestion could be investigated. "The analysis of covariance may not help in this case since it is sensible to non-random assignment in the presence of fallible covariates as well as to nonlinear regression, where there are large initial differences in the groups" (p. 211). Some other points made by Wiley are: The multivariate test statistics and their associated probability levels are not used. The adjusted means are not reported for the analyses of covariance. Tests of homogeneity of regression do not precede the analyses of covariance.

However, Wiley inclines towards the positive and mentions a number of commendable features of the study, among them "..... the monitoring of the treatment effects which allowed rather more precise definition of the various strategy-laboratory combinations. This is especially useful for those who wish to base decisions on the study" (pp 211 - 212). It is noteworthy that this point, like so many others, has been quite differently commented on by competent reviewers.

In the October issue, 1969, of the *Modern Language Journal*, the Pennsylvania project was fiercely criticized in a number of articles. Some of them were very negative in tone, and one wonders whether the authors had an axe to grind. Anyway, there is reason to believe that at least some objectivity was sacrificed in the heat of argument. We shall be brief in our comments.

Hocking, concentrating on the comparisons between laboratory systems, seems to be accusing the project staff of sabotage as far as the language laboratory side was concerned. Hocking seems to advocate more restricted projects than the Pennsylvania study which he thinks involved too many inponderables and uncontrolled variables. However true this may be, a strong need was obviously felt in the mid-1960's that a study of this dimension should be undertaken.

Clark's main criticisms (p. 388 ff) include: non-random assignment of classes to methods, no clear distinction between methods, faulty scales for controlling teacher adherence to strategy; all these items have appeared above. However, Clark's argument on p. 394 has a strong resemblance with our own discussion of "Hypothetical Treatment Effects" (see p. 22 below): "Within the Pennsylvania Project, the most powerful demonstration of superior pedagogical efficiency for one or another of the three teaching methods would have been for that method to satisfy all of the following conditions: 1) to prove superior for both the

French and German groups rather than for a single group; 2) to show superiority on all three measurement occasions (first- and second-year tests for the original group; first year test for the replication group); 3) to show similar results for closely related tests, as within a single skill area; and 4) to prove superior to both of the other two methods, rather than to only one of these methods. To the extent that these outcomes are not reflected in project results, it becomes necessary to introduce explanatory hypotheses which may become so diverse and complex as to reduce considerably the possibility of identifying a single factor - such as inherent superiority of a particular teaching method - which would account for the observed results". Clark contends that the only safe generalization that can be made for the results of the study is that the majority of comparisons show non-significant differences among the teaching methods. However, he does not accept this as evidence of the pedagogical equivalence of the methods but considers the possibility that true differences may have been concealed by uncontrolled factors.

Otto's review (p. 411 ff) is primarily focused on the area of teacher activities within the project. He contends that the MLA Proficiency Tests do not measure pedagogical proficiency, that several teachers were assigned to teaching strategies against their preference, that assignments were not based on effective screening techniques (which would have helped the project personnel to determine if the teachers had the ability and experience to follow a particular strategy), that the so-called orientation sessions for teachers did not provide exemplary models of effective teaching behaviours for each strategy, that the orientation sessions were no work-shop sessions (which was what was needed), that assistance and supervision was not sufficiently provided, that the *Teacher's Manual* was poorly organized. In short, Otto is strongly negative towards the project, at least those aspects of it which regard the teachers and the part they played.

Valette, in her review (p. 396 ff), mentions one feature which most reviewers have touched on, namely the fact that the complex findings of the Pennsylvania project have been over-simplified and misinterpreted in various press releases. Stressing the disservice such journalism does to both the project personnel and the foreign language teaching profession as a whole, she urges anyone really interested in the results to read the full reports.

One interesting comment by Valette is the following (p. 397): "(Consequently), the section of the Pennsylvania Project which contrasts teaching approaches has almost become out-dated before the results have been disseminated". Her argument is that, in 1969, the distinction between "traditional" and "audio-lingual" is losing some of its relevance because the new traditional texts (the "third generation" texts, in Valette's terminology), make creative use of dialogues and pattern drills whereas (the "second generation") audio-lingual texts give attention to formal grammar. This phenomenon has an obvious resemblance with "the struggle towards the middle", which was discussed in our previous report (Levin, 1969, p. 79).

Some of Valette's criticisms of the study are the same as those discussed above, some may be new: TLM students received more contact with the spoken language than was intended, the contents of the Cooperative tests favoured TLM students (TLM students did much more poorly on this test, however, than one would have anticipated), the criterion test was too difficult, the student opinion scale is dubious (an expert on attitude testing ought to have evaluated the instrument), etc.

Her main point on the use of the language laboratory is that, in the lab, one tape was played to the entire class; thus the lab was not used for individualization. "..... we must distinguish between the physical installation which we term a language laboratory and the use we make of that laboratory" (p. 404).

Finally, mention should be made of Valette's proposition that, in modern languages, criterion-referenced tests should be developed. According to her, the Pennsylvania project had specified "expected levels of proficiency" but had no tests available to assess whether the pupils reached those levels.

The last review in the "October issue, 1969" that we shall comment on is that of Aleamoni and Spencer (p. 421 ff), who are very critical: "In general, the objectives of the study are stated more broadly than the study seems capable of handling; and it covers areas so diverse that it would be difficult for any study to accomplish them" (422).

The authors criticize the project for being unwieldy and unmanageable.

Furthermore, the project staff is accused of being subjective and biassed in planning the study: "Many of the statements in the early

pages of the reports are statements of belief, opinion, or attitude, which set the stage for the research design. These statements appear in the reports without evidence or documentation " (p. 423). Some of the more specific criticisms concern the (alleged) misuse of the interest, attitude, motivation and teacher factor scales, the decision not to include students for whom complete data were not available, use of the same test as both a covariate and a criterion when the covariate had been subject to the effects of the treatment, etc. Of all the recommendations to the teaching profession, forwarded by the project staff at the end of the reports, none seem to escape Aleamoni's and Spencer's criticism.

Later on Dr. Smith wrote a reply to the October, 1969, *Modern Language Journal* (Smith, 1969 c). When he states that "Some reactions have been of the highest professional quality, some reflect simply a lack of understanding, others smack of panic" (p. 3), he refers to all reviews until that date. Concerning the specific MLJ reviews he contends that they "often present a distorted view of the Pennsylvania Studies in that they suffer from (1) a narrow and insulated viewpoint; (2) overt hindsight; (3) personal interpretation; (4) inconsistency; and (5) obvious oversight. This is tragic, especially in that the *Modern Language Journal* attempts to be a responsible professional journal but will not protect its contributors nor its readers from obvious oversight, choosing to let errors stand as definitive statements of the research" (pp.5-6). For some reason, the reviewers had had no contact with the project staff, which might have led to a correction of errors - if there were such - or at least to a relaxed atmosphere, more advantageous to scientific cooperation.

Dr. Smith points at a number of issues where the reviewers have different, not to say opposed, opinions. However, we shall not discuss his counter-arguments here, nor try to pass any kind of value judgment on them. It seems a difficult task to make a reliable and comprehensive evaluation of the Pennsylvania project in all its complexity. At any rate, the contrasting views of competent researchers on various aspects of the project, is one indication of this.

Whatever significance the project results will have in the long run, the following statement may be made with confidence: being contrary to the expectations of many foreign language teachers, the project results have initiated a debate that will in turn initiate wholesome rethinking on various aspects of foreign language teaching methodology.

EXPERIMENTATION IN A FIELD SETTING - SOME REFLEXIONS

Comparative Experiments - Pros and Cons.

The present study is a case of variable-manipulating, comparative experimentation in a field setting. Since the general value of such research has occasionally been questioned, a comment may be appropriate.

A classic in this debate is Scriven's (1968) article, where the principles of formative and summative evaluation are introduced and, which is of greater interest here, where Cronbach's (1963) "despair over comparative studies" is optimistically contradicted. "If we have really satisfied ourselves that we are using good tests of the main criterion variable (and we surely can manage that, with care) then to discover parity of performance *is* to have discovered something extremely informative. 'No difference' is not 'no knowledge'" (Scriven, p. 67). Scriven apparently holds the view that the comparative field study has a definite (though by no means unlimited) place in evaluation.

A representative of the negative attitude towards field experimentation is Grittner (1968) who, when commenting on the bulk of studies presented by Stephens (1967), concludes: "In short, half a century of such 'research' has told us almost nothing about the relative superiority of one educational strategy over another!" (Examples of the areas which Stephens reported on are the following: large vs. small schools; large vs. small class size; accredited vs. non-accredited teachers; progressive vs. traditional education; live teachers vs. TV; lecture method vs. discussion method; team teaching vs. traditional teaching; and homogeneous vs. heterogeneous grouping of students). "Tables showing standard deviations, covariance, F-ratios and the like are very impressive; however, if the ultimate result of such studies is that they cancel one another out, perhaps we should ask for a cease fire while we search for a more productive means of investigation" (p. 7).

Wiley (1969) makes a distinction between conclusion- and decision-oriented research. The former is performed so that the investigator may draw conclusions about the phenomenon he is studying. Conclusions,

however, are tentative by nature and may be modified as more evidence is accumulated. Decision-oriented research, on the other hand, is performed to gather evidence which will be used for generating decisions about actions to be taken. Wiley gives the example of a school superintendent who cannot wait for twenty-five years of accumulated evidence before deciding whether to purchase a language laboratory. If he does so, he will really have decided against it (p. 209). Wiley further argues that the concern for the quality of evidence must be greater in the case of decision-oriented research; decision-makers cannot wait for ambiguities to be clarified by subsequent investigations. Under these circumstances, the methodology of research becomes extraoridinarily important.

The point that we want to make here is that Wiley seems to come rather close to the traditional design proposed by Campbell and Stanley (in Gage, 1963) when suggesting proper evaluation methodology. The main difference seems to be Wiley's greater concern with the criterion tests to be used in program evaluation ("It is not individuals among whom we wish to discriminate; rather it is programs", p. 208). His philosophy of evaluation thus seems to be quite similar to Scriven's. In spite of the difficulty of constructing reliable evaluation instruments, Wiley seems to be in favour of experimentation in school settings.

Stanley (1970) regrets the present state of affairs in educational research, which, according to her, is characterized by the paucity of controlled experimentation. "Apparently there is more lack of intent, money and technical resources than of available, applicable methodology. Those critics of experimentation for evaluation who say that controlled, variable-manipulating experimentation may be splendid for stands of alfalfa and weights of pigsbut inapplicable to education do not adequately appreciate the generality of Fisherian and neo-Fisherian methods. ....Inflexibility is more in the minds of planners, researchers, and critics than in the methodology itself. Of course, there is no royal road to new knowledge; it is not easy to experiment with human beings, whether they are medical patients or school pupils. In my opinion, however, controlled experimentation and some quasi-experimental designs are important methodological tools of the education evaluator. Recent attempts to rule experimentation inapplicable because other methods are also useful seem misguided" (p. 107).

The survey of opinions for and against experimentation in the natural school setting might have been made more extensive. For the moment, however, we shall be content with this list of contrastive views. Textbook writers in the branch of educational research often present an almost overwhelming list of difficulties in experimentation but end up with words of encouragement, urging the student to use experimental methods whenever they are feasible.

Let us conclude this section by quoting Wiley once more (ibid, p. 210): "In any research study, especially one conducted in a field setting, it is impossible to do everything 'right'. There are always going to be unanticipated contingencies and contingencies which, although anticipated, are practically (usually monetarily or cooperationally) impossible to avoid. The main goal is to spend the most time, effort, and money to avoid the most 'important' pitfalls to the validity of the findings and their interpretation. One problem is that the 'importance' or relevance of each pitfall is different for different individuals".

The GUME Project - Some Comments.

In one of the earlier GUME reports (Levin, 1969, p. 27 ff) our first three studies were discussed in relation to Carroll's chapter "Research on Teaching Foreign Languages" in Gage's Handbook (Gage, 1963, p. 1060 ff). Here we shall avoid unnecessary repetition; however, a few points will be made.

In GUME 5, as in the first four projects, we do not have the advantage of what Carroll calls a natural zero-point in second-language acquisition. The experimental population consists of pupils in their fifth year of English, as compared to the fourth year in the previous studies. Although prior knowledge in English is controlled statistically by analysis of covariance (to the extent that our Criterion test measures this), it is obvious that the *amount* of treatment (teaching) must be large for *differences* between the various treatments to appear. Concerning differences, hypothesized or real, there is one question of great concern: Should one use radically different treatments, thereby increasing the chances for a "positive" outcome but decreasing the external validity of the findings, or should one construct different but "realistic" methods that might be used later

at school, thereby decreasing the probability of obtaining "positive" results? Posing the problem in this manner is perhaps somewhat naïve, but it has to be solved, anyway. For GUME 5 it was decided to have realistic methods tested in the classrooms at the risk of not obtaining any results rather than trying to get results with drastically dissimilar methods and then be left with the question of how to interpret these results and what use they could be put to.

Another circumstance decreasing the probability of obtaining positive results is the fact, not peculiar to GUME but rather general, that pupils vary in a number of aspects, and that *this variation is treated as error in the analyses*. Incidentally Carroll (1969, pp 233-34), when reviewing the Pennsylvania Study, notes that "another unassailable fact arising from the study - and one that carries at least some surprise - is that *classes* vary enormously in average performance". Without anticipating our results we may perhaps state that the same observation was made in the present study; the differences between the school classes, let alone between the individual pupils, was enormous. Hopefully a good deal of this variation is held constant in the analyses of covariance, but it would be a false assumption to believe that all that variation, for instance in Post-test scores (an indication of a corresponding variation in general ability, motivation, reading facilities in the home, day-dreaming tendencies and what not) could ever be held constant, experimentally or statistically.

Hypothetical Treatment Effects. x)

The present investigation implies a comparison between three teaching strategies. No assumptions are made about the superiority of any one method; to use a different terminology, the null hypothesis is being tested. The experimental design should be such as to make interpretations of the results as clearcut as possible. Of all the theoretically possible outcomes, some are more difficult to interpret than others. In this section we will briefly discuss specific interpretation problems that may arise.

The three teaching strategies being compared are

Im Ee Es

x) This section is identical with the one in Levin (1969, p. 29 ff).

On the one hand the effect of explanations is compared with the effect of non-explanations, on the other one method utilizing the source language (Swedish) is compared with two methods utilizing the target language (English). An ideal design for isolating the effects of explanations/nonexplanations, source language/target language would have to include an Im_s, i.e. Im-Swedish, variant. However, since such a method is impossible by definition, and, accordingly, could not be included in the design, the interpretation problems indicated above will arise in certain cases.

When comparing three strategies, the following main results are possible:

a) two methods equal and better than the third (3 possibilities)
b) one method better than the two others, they being equal (3 possibilities)
c) method X better than method Y better than method Z (6 possibilities)
d) the three methods equal.

According to a) above, the following three outcomes are possible in the GUME project:

1. Ee = Es > Im
2. Im = Ee > Es
3. Im = Es > Ee (?)

In case 1 the facilitative learning effect is unequivocally due to the explanations, in case 2 to the use of English, whereas in case 3 the result could not be logically explained. The superiority of methods Im and Es can be accounted for neither by reference to language of instruction nor by explanations.

Correspondingly there are three possible outcomes according to b) above.

4. Im > Ee = Es
5. Es > Ee = Im
6. Ee > Im = Es (?)

In case 4 the non-explanation method is unequivocally better than the two explanation methods, in case 5 the facilitative effect can be traced to the use of the source language, whereas in case 6 the outcome is impossible to interpret. According to c) above, six results, approximately identical to the six just presented, are theoretically

possible. Our intention here is only to predict difficulties of interpretation in general, and we will not discuss interpretation problems under c) further. Concerning d) (the three methods equal) it should be remembered that such an outcome does not *prove* that there exist no differences between the methods (as is well known it is a logical impossibility to prove the null hypothesis). One possible explanation might be that the experiment, as it was planned and executed, did not succeed in detecting actually existing differences between the methods.

To sum up:

The experiment makes possible comparisons between three methods of instruction. Theoretically thirteen different outcomes are possible. Some of them would be impossible to explain, or rather, would arouse doubts about the experiment, notably the experimental control of the three teaching strategies. We may have good reason for returning to the interpretation problem in the results section.

METHODS IN FOREIGN LANGUAGE TEACHING

Since the present project is an effort to assess the relative effect of different teaching approaches, it is of the utmost importance to have a clear definition of the methods used. In the following discussion of some current methods special attention should be given to principles concerning the teaching of grammar since the particular problem of this study bears upon the teaching of grammatical points.

Current Methods and Their Historical Background.

The prevalent methods used today in foreign language instruction are derived from two different schools of thinking, that is, those who maintain that language is a mechanical process and those who think that it is an intellectual process. The grammar-translation, the traditional method as well as the cognitive code-learning theory belong to the second category while the direct method, the oral approach, and the audio-lingual habit theory belong to the first one.

The direct method was created as a protest against the grammar-translation method used in the schools towards the end of last century. In the direct method the mother-tongue of the pupil is ruled out in the instruction and understanding of the new language is arrived at by demonstration. Grammar is to be learnt inductively just as when a child learns his mother-tongue, and listening to and speaking the language are primary to reading and writing it. The great linguist Jespersen is one of the advocates of the direct method. He described this method exhaustively in *How to Teach a Foreign Language* (1904). He was also one of those who, at the Philological Congress at Stockholm in 1886, were behind the Quousque Tandem movement in Scandinavia and spoke for a reformed foreign language instruction. He rightly pointed out that it is difficult to name any single person who has formulated the principles of the direct method, even if the names of Berlitz and Gouin come to our minds when we think of the first days of the direct method (p. 3). In fact, Berlitz schools for languages exist all over the world to this day and they advertise "total immersion" courses where the foreign language is spoken by pupils and teachers from the very first lesson.

In Sweden, however, the direct method did not conquer the field even if it gave rise to a heated and animated discussion. It also caused two monolingual grammars to be written, namely those by Karin Åhlström, *Engelsk Språklära* (Stockholm, 1894) and Daniel Elfstrand, *An English Grammar for the Use of Swedish Schools* (Stockholm, 1897). They were not, however, followed by any others until 60 years later. The debate nevertheless caused a compromise between the old and the new method in the official recommendations for schools. Thus *Kommittébetänkandet* of 1902 laid down that the pupil in the grammar-school was to learn the accidence and the syntax of the foreign language, though inductively. A grammar-book was to be used but the rules should be studied in connection with the reading of texts (p. 132).

The grammar-translation method or the traditional method of last century with its emphasis on the disciplining of the pupil's intellect by grammatical puzzles and the testing of his knowledge by translation as predominant features is very different from the version used today. It is true that the main emphasis of this method is still on "conscious control" of grammatical concepts. The study of grammar should be deductive. However, the spoken aspect of the foreign language is not completely set aside today. Thus both the direct method and the traditional method exist in modified versions nowadays. Besides, they may mean very different things to different members of the profession.

Since the 1940's the application of structural linguistics and behavioral psychology to teaching techniques has caused a new orientation in instructional theories in the U.S.A. The new principles of the oral or audio-lingual method can be summarized in the slogans: "Language is speech, not writing", "A language is a set of habits", and "Teach the language, not about the language" (Moulton, 1961, pp. 86-89). How grammar was looked upon by the proponents of these two methods will be further clarified in the following quotations:

"If by 'grammar' we mean any of these things - the memorizing of paradigms, or the logical analysis of sentences, or the learning of the rules of a philosophical or universal grammar, then we can easily agree that we must approach a new language by a more 'natural' method. But "grammar" from the point of view of modern linguistic science means something entirely different from any of the matters enumerated above and it can be used in a manner that does not in the least conflict

with the so-called natural way in which a child develops in the grasp of his native language" (Fries, 1946, pp. 27-28).

"Talking about the language is a fascinating activity at any stage, but if it is done in the classroom and in the students' mother tongue, it is a waste of time. Nothing can replace the practice required for mastering the language. Only when the production of the foreign language has become automatic can the student be said to have made any real progress. Conscious analysis can only slow down the process for the speaker" (Northeast Conference, 1958, p. 38).

"Analysis is important in its proper sphere, but analogy is used instead through pattern practice to produce a control of language structure without the time and effort required for grammatical explanations.... Since every speaking person has mastered his own language through imitation and analogy without benefit of analysis it stands to reason that something of this ability will aid him in the learning of another language. Pattern practice permits this ability to function" (Nelson Brooks, 1964, pp. 146-147).

A return to the mentalistic orientation has recently taken place in the approach called the cognitive code-learning theory (Carroll, 1965, p. 273 ff). Its background can be found in the psycholinguistic science and also in Chomsky's theories on language acquisition (Chastain, 1969, pp. 97-106). This approach puts primary emphasis on the student's comprehension of structure. It questions the excessive use of patterned drills but does not go so far as to maintain that the student can arrive at fluency in the foreign language without applying the rules of the foreign language in various situations which by the way are similar to the pattern practice techniques.

The Authorized Curricula for Schools.

Until the autumn term of 1970 Läroplan för Grundskolan (the Authorized Curriculum for Schools, 1962, henceforward referred to as Lgr 62) was still in force. Since then, however, it is being replaced by Läroplan för grundskolan (1969) with its Supplement in English (abbreviated Lgr 69 II:Eng). It will be in effect in grade 8 when the autumn term starts in 1971. In Lgr 62 grammatical knowledge of the foreign language was regarded as a means to an end. The pupils should not be burdened by unnecessary analyses and rules but learn the grammatical structures by systematic drills of different kinds. The study of grammar should

be cyclical. If a rule or an explanation is deemed necessary, the inductive procedure is said to develop the pupils' power of observation. It is, according to Lgr 62, advisable to use the Swedish language when grammar is being discussed, if no real clarification can otherwise be attained. Before the rule is formulated, the pupils should have heard several examples of the pattern in question, and a visual representation should also have been given so that the pupils are fully aware of what they are practising. Oral translation from Swedish into the target languageis not excluded when practising grammatical points (pp. 197-8).

Lgr 1962 can be characterized as favouring an approach which has much in common with a modified direct method, but the Supplement of 1969 has a still clearer tendency towards the principles belonging to the mechanistic school of language acquisition. Lgr 69 II:Eng does not mention translation into the foreign language as a means of promoting a functional control of grammar. The insight which the pupils should acquire about the structure of the target language is said to be arrived at first and foremost by systematic drilling. There should be at least ten (sic!) examples of the pattern in question on each instructional occasion. Overlearning is considered necessary for a lasting command of the language. Lgr 69 II:Eng still allows rules and comments on the language, but they are to be the final step in the teaching situation. Furthermore, formulations of rules should describe exclusively the structure of the English language. If the Swedish language is used for observations on grammar, which, according to the Recommendations, is permissible in rare cases, no comparisons with Swedish usage should be made (pp. 12-14).

Towards a Definition of the GUME Methods.

In the GUME project the word "method" does not signify an entire teaching strategy, which rightly should include many more matters than the teaching of grammar. For that reason none of the existing denominations of methods were used in this project. Instead the terms *Implicit* and *Explicit* were chosen so that all identification with earlier known teaching techniques was avoided. The theories behind methods in general use, however, can be applied to the GUME methods Implicit and Explicit, that is to say that these two methods can be pidgeonholed into the two theories of whether language acquisition is

a mechanical or a mentalistic process. The Implicit and Explicit methods have been modified by discussions and experience and do not represent any unrealistic or theoretical creations. On the contrary they are instructional styles which can be used in ordinary classrooms and which are no doubt practised in the everyday teaching of foreign languages in Swedish schools.

The Implicit method. The adjective 'implicit' can mean that something is tacitly understood or implied as opposed to expressly stated (*The Random House Dictionary*, 1968, p. 667). Thus the Implicit method here stands for an approach where systematic series of drills are to result in a subconscious assimilation of the rules. The learner's attention is directed to the crucial features of the sentence by way of analogy or contrast. No verbalized explanation or generalization about the language ever occurs within this method. Nor is the Swedish language used on any occasion.

It is clear that in the exclusive use of the target language the Implicit method has a facet in common with the original direct method. It is, however, also evident that it owes the heavily-structured drills as well as the dialogues to the audio-lingual method. Proponents of unmodified versions of these two methods also agree that generalizations about what happens in the sentence are by no means necessary. The Implicit method is thus bound by close ties to the mechanistic school of thinking.

The Explicit methods. In this project the explicit method represent the school of thinking that maintains that the acquisition of language is an intellectual process. The students within these methods are made consciously aware of the functioning of the language by verbalized generalizations and explanations about what they have just heard, spoken, read, or written. The grammatical point to be learnt is at first clearly brought into focus in dialogues and patterns and then commented on in a way considered suitable for the age of the learner and his level of knowledge. The explanations, or perhaps rather the generalizations, concentrate more on descriptive observations on how the parts of the sentence behave than on why they behave like that.

In the teaching of grammar the terms "deductive" and "inductive"

are sometimes used. If the deductive process is used, the learner is given the rules before drilling takes place. The term "inductive", however, can apparently indicate more than one line of action. Sometimes it seems to mean that the rule is inferred by the student from the learning materials, but not given an explicit formulation (cf. the original direct method p. 22). In other cases it implies that the students induce the rule on their own and help give it an explicit formulation, a procedure mentioned, for instance, in Lgr 62 and Lgr 69 II:Eng. Furthermore, an inductive procedure can mean that the pupils receive grammatical explanations after what is called "functional practice" (Kiat-Boey Lim, 1968, p. 10). The teaching of grammatical points within the explicit methods is inductive in the sense that the pupils have in most cases seen quite a few examples before they are given an explanation. It goes without saying that with "canned" lessons there can be no trial-and-error experimentation on the part of the pupils when it is the question of formulating the explanations.

The project had two variations of the explicit methods. The first version, the Explicit-English method, gave the explanations in English. The second, the Explicit-Swedish variety,used the Swedish language. The explanations in English and Swedish are, however, not merely translations of each other, as the Swedish version also includes comparisons with the corresponding Swedish structures. This is quite in line with the recommendations of Lgr 62, but as pointed out earlier (p. 25) it goes against the directions of Lgr 69 II:Eng.

Stages in Foreign Language Learning.

A further characterization of the features of the three GUME-methods can be brought about by applying to them the definition of methods given in a work paper (The Center for Curriculum Development, Philadelphia, USA) on stages in foreign language learning:

"Second language learning, no matter through which approach or method, follows a sequence of four basic steps: presentation of the item, explanation of the item, repetition to mastery, and transfer to appropriate real-life situations. It is in ordering, emphasis and style of these four steps that 'methods' differ. Every approach contains these either implicitly or explicitly in some arrangement and all materials are designed to contribute to one or more of the basic steps".

If the above definition of methods is applied to the present study it is found that the ordering of steps is presentation at the beginning and transfer at the end of the four stages. Repetition and explanation could by turns occupy the second and third places.

Repetition is said in the above mentioned work paper to indicate all exercises, patterned as well as of a freer kind, and oral as well as written. The present writers would object to the expression "repetition to mastery" about this project, however. The presumption has never existed that the experimental instruction (6 lessons) would lead to complete mastery of the structures practised so that mistakes were excluded in future and much less that it could be predicted when this mastery of the structures in question was to occur. Even in an individually arranged instructional situation the possibility of stating when a grammatical point has become part of the learner's competence is extremely problematic, and the situation is infinitely more complicated when all pupils are taught in the same way in a classroom.

Explanation as discussed in the work paper stands for either a verbalized generalization or an arrangement of the learning materials in such a way that the student would infer the rule for himself. The latter procedure is what is supposed to take place in the Im-approach. It should be noticed, though, that when the place of the explanations is discussed in this report, it is the question of the explicit explanations only.

Methods differ, according to the work-paper, not only in the ordering of these four steps, but also in emphasis and style. The grammar-translation method must of course give great attention to explanations while, on the other hand, the audio-lingual method concentrates on the remaining three items. GUME 5 had explanations in the two explicit strategies, but they took up very little time compared to the emphasis given to repetition and transfer exercises. Presentation took up more time in the earlier lessons than in the later ones. Finally, the style of the approach was influenced by the fact that the teacher variable had almost completely been eliminated. The oral and written drills were therefore very often heavily structured so that the students should be able to go on with their work without individual help from the teacher. This also meant a certain rigidity in the approach, but it could not be avoided in the circumstances.

Debatable points. What is causing great disagreement among language teachers at present is the place, emphasis, and nature of the explanations. Ian Dunlop, for instance, writes that the patterns of grammar should be pointed out as well as practised. The ideal hierarchy between them should be: Example + rule + examples (Dunlop, 1970, p. 43).

As mentioned earlier Lgr 69 II:Eng stresses that explanations, when they are considered necessary, should be introduced as the last step so that the four basic stages should be: presentation, repetition, transfer, explanation (p. 13). Finally, Ausubel, writing on teaching grammar to anybody but young children, holds that deductive use of grammatical generalizations is decidedly more efficient than discovery learning. If the learner is given the rule at the start both the generalization and the experience of applying it in appropriate cases are transferable from the beginning of the exercise (Ausubel, 1964, p. 422).

It seems highly improbable to the present writers that it is possible to single out any one procedure when teaching grammatical points and declare it to be the one saving path for all stages, ages, and different kinds of learners. A more profitable line of action is the attitude to various methods taken long ago by H.E. Palmer, who advocates a multiple line of approach (1922, pp. 108-15). The same flexibility on the part of the teacher is recommended by Rivers in her description of the eclectic method. In this "method" the best techniques of all well-known language methods are used when the teacher thinks it is to the purpose (1968, p. 21).

Recent Methodological Discussion.

Experiments trying to assess the relative effectiveness of various teaching techniques have as a rule only demonstrated the futility of the efforts to find clear-cut and statistically significant differences between the methods.A case in point is the result of the Pennsylvania study, which has been discussed above at some length (see p. 6 ff).

The lack of success in the evaluation of methods is usually attributed to the great number of variables involved in the instructional situation. This is, for instance, an attitude taken by Bosco and Di Pietro in a recently published article (1970, pp. 1-19), but in addition they stress the fact that many facets are common to more than one method and consequently that the result when entire strategies are compared can only be deceptive if there is any result at all. Instead they suggest the use of a framework on linguistic and psychological bases where the distinctive features of each strategy are analyzed. They think that not only can a strategy be clearly defined in this way but it may also lead to the definition of an optimal instructional strategy.

Similar ideas but not the same optimism is expressed in a booklet published by the Centre for Information on Language Teaching (CILT, Reports and Papers 2, September, 1969, pp. 27-33). The last two chapters stress the necessity of isolating the different factors which make up the methods to be assessed. Even if this can be meticulously done, it is considered that with matched groups and matched teachers one can get no further than establishing that one course might suit, for instance, teachers who are fluent in French while another might be better for teachers who are not. Language teaching efficiency is a joint product of materials, techniques and the ability and personality of the teacher in unknown proportions. It is suggested that the teacher is the most important factor in most classrooms (pp. 31-2).

Doubt as to the value of assessments of instructional strategies has also been forwarded on other grounds. Professor Bruce Pattison of London University has expressed the opinion that evaluations of different methods are both impossible and unnecessary. According to him they are impossible because the teaching situation involves so many variables that an effective control of them is out of the

question and unnecessary because it is already known to language teachers what good teaching should be like (Pattison, 1970, pp. 3-10).

The difficulty of establishing in an experimental way which instructional method is the most effective has not stopped theories from being formed on what constitutes successful foreign language teaching. The grammar-translation method was heavily attacked in the U.S.A. when the audio-lingual method was introduced. An example of this is Bloomfield's description of the state of affairs in the U.S.A.: "The large part of the work of high schools and colleges that has been devoted to foreign-language study, includes an apalling waste of effort: not one pupil in a hundred learns to speak and understand, or even to read a foreign language" (Bloomfield, 1933, p. 503). Similarly, the audio-lingual method is today accused of not keeping abreast of recent developments in linguistics and psychology. It is now being challenged by the cognitive code-learning theory which opposes the audio-lingual habit theory on two fundamental points, that is, what language is and how it is acquired (Chastain, 1969, p. 105). The transformationalists' views on language acquisition have not yet affected foreign language teaching to any appreciable extent and it is difficult to foresee what impact it will have on language teaching in future. Their theory implies that the student should make observations on the deep structure of the foreign language. Surface structure similarities as demonstrated in pattern drill activities is according to this theory completely unenlightening. The burden of language acquisition is placed with the learner while much less stress is put on the conditioning contingencies of his linguistic environment (Jacobovits,1968, pp. 90, 106).

The criticism directed against the audio-lingual habit theory has not remained unanswered. James W. Ney writes that the audio-lingual or oral method has in reality undergone changes since the 1940's and 1950's in accordance with the linguistic climate (Ney, 1968, pp. 3-13). It now stresses the necessity for the learning material to be meaningful and realizes that drills should be visually supported and not only aural. The transformationalist opinion that "conditioning" plays no part in the language learning process is refuted by Ney. To prove his point Ney compares the two sentences "I am smart, aren't I" and "I am smart, am I not" as uttered by native children. The deep structure is the same in both cases, and the surface structure is

different only because of different "conditioning" processes in the children´s linguistic environment. "Conditioning" is, according to Ney, a very important ingredient in language learning, and hence he opposes the transformationalist view that pattern drills are meaningless activities in the language acquisition process. He also repels the notion that the audio-lingual habit theory should exclude understanding of the functioning of the language and quotes Fries´s words: "Generalizations concerning structure, or grammar, are a regular feature of the 'oral approach' although they are always intimately related to the oral practice of the language" (1946, p. 7). While Ney thus sees the need for revision and inprovement of the audio-lingual method, he still thinks that there is no better alternative.

In Sweden the lively discussion carried on in daily newspapers and professional publications since 1968 on the alleged superiority of in turns the direct method and the traditional or grammar-translation method culminated during the spring of 1970 in an address signed by more than 2.000 language teachers at the gymnasium level and handed over to the Minister of Education. In it the teachers stated that the results of the foreign language instruction had deteriorated rapidly during the last years. They blamed the situation on the monolingual instruction recommended in the Authorized Curriculum for the Gymnasium (Läroplan för gymnasiet, 1965). In ten points they made it clear what changes they wanted in future. They desired the prescribed methodology to include features from the traditional method as well as from methods created more recently. Not only should teachers in foreign languages but also those in Swedish endeavour to give the pupils the grammatical insight appropriate to the different age-groups. The grammarbook should partly build on contrastive analysis and the rules should be in Swedish. The oral instruction should be sufficiently backed up by written exercises, and translation from and into the foreign language should be used as an instructional means whenever it was considered to be to the purpose.

In the autumn a commented summing-up of the debate was published (Edwardsson, 1970). Noteworthy among the numerous contributions is an article by Arne Klum where he points to the undifferentiated classes and the complete lack of instruction in grammatical terminology as the villain of the piece. In such conditions methods are, in his opinion, insignificant as causes of bad results. It could also be added that

no motivation for theoretical studies and bad disciplinary conditions would preclude the success of the best of methods. Most of the contributors in Edwardsson´s book, however, put the blame for the alleged catastrophic results of the foreign language instruction on the so called modified direct method officially recommended (p. 40). In reality, though, the Authorized Curricula give the teachers a fairly wide choice of instructional style and while stressing the functional aspect of language acquisition, they do not forbid grammatical analysis. Officially it has been acknowledged that the results of the foreign language instruction may not today be of the same high quality as a few years ago, but the widening recruiting to the upper school forms is assigned as the reason of this (p. 71).

To an unbiassed observer of the situation in Sweden it must be clear that there can be no single reason, applicable to all stages and age-groups, of why the result of the foreign language teaching does not reach the desired quality. That the methodology has been so generally picked out as the scapegoat could be because this is a field where even earlier opinions differed most widely and, witness Edwardsson´s book, most passionately.

A DESCRIPTION OF THE LESSON SERIES

Lesson Design.

The lesson materials of GUME 5 consisted of speaking, writing, and reading modules, but it was not a matter of course, as in the GUME 3 experiment that the order between these activities should be: 1. Speaking. 2. Writing, and 3. Reading. So, for instance, writing drills can occur both at the beginning and at the end of a lesson. The exercises were the same for all the three strategies with the exception that the Im-group had a little more practice during the time taken up by explanations in the other two groups.

Below follows a table showing the distribution of time for the totality of the instruction in the different groups as well as for the three activities speaking, writing and reading:

Table 1: Allocation of Time within the three Strategies.

Total time of instruction:	Im:	2 hours	59 min.	19 sec.
	Ee:	3 "-	4 "-	51 "-
	Es:	3 "-	2 "-	2 "-
Total time of speaking in the	Im-group:		83 min.	47 sec.
	Ee-group:		68 min.	- sec.
	Es-group:		65 min.	17 sec.
Total time of writing in the	Im-group:		66 min.	20 sec.
	Ee-group:		60 min.	17 sec.
	Es-group:		61 min.	4 sec.
Total time for the reading in the	Im-group:		27 min.	35 sec.
	Ee-group:		27 min.	24 sec.
	Es-group:		28 min.	21 sec.

The time taken up by the six lessons in the different strategies is nearly identical, and the two activities speaking and writing seem to be comparable in time as far as the Ee and Es groups are concerned. It is also evident that the speaking drills in the Im-group benefited

more than the other activities by the time given to explanations in the other groups. Finally, the time for the reading of texts is fairly permanent in all the three groups.

Listening is included in all activities, but it has only been allowed to occupy a separate column when no special drill had been built upon a listening comprehension exercise (see table 2, next page).

Written activities. Speaking is a skill that is very much stressed in the Recommendations of the Supplement of Lgr 69. Writing, on the other hand, does not receive the same generous treatment either in emphasis or in columns in the Supplement. So, for instance, it is found there that the time allotted to writing can be cut down to a bare minimum in ak while the oral instruction should dominate at all stages within "grundskolan" (p. 33). As mentioned before, the allocation of time to the various activities was not streamlined in advance, and no definite plan existed which said that writing should get as much or less time as speaking. It was however borne in mind that both pupils and teachers complained during the first experiment about the shortage of time for the written exercises which seldom allowed the pupils to finish anything.

The special conditions under which this experiment has been conducted made some activities easier than others to carry through. Exercises in writing belong to the activities which are easy to start from a tape-recorder. Free conversation is, on the contrary, fairly difficult to have under control from a tape just as work in groups. All these circumstances have contributed to giving the writing drills as much time as the speaking drills. It should be noted that if Lgr 69 II Eng does not put an equal stress on the two activities, this is a fact which is of little relevance to the present study.

Oral drills. It is evident that in GUME 5 all the three strategies have taken from the audio-lingual method the dialogues in which the pupils perform one of the parts during the exercises. In accordance with the principles of the audio-lingual method the stimulus in these drills was mostly aural (in a few cases pictures were used). Only during a few repetition exercises did the students see the drills in print. Of late criticism has been directed against this system of presenting pattern drills in audio-oral form only. Wilga Rivers points

Table 2 : Allocation of Time per Lesson within the Three Strategies.

		Introduction	Speaking	Writing	Reading	Listening	Explanations	Total
1.	Im	0.45	10.26	8.45	2.50	0.52		23.38
	Ee	0.45	8.45	8.19	2.44	0.58	2.38	24.09
	Es	0.40	7.55	8.33	2.52	1.00	3.13	24.13
2.	Im		21.13	5.17	4.55			31.25
	Ee		18.00	5.30	4.51		4.18	32.39
	Es		17.37	5.37	5.04		4.05	32.23
3.	Im		10.15	10.20	6.30			27.05
	Ee		6.59	7.41	6.33	1.15	6.17	28.45
	Es		7.15	7.58	6.30	1.09	5.40	28.32
4.	Im		16.43	12.13	2.57			31.53
	Ee		13.07	10.38	3.01		6.00	32.46
	Es		11.15	11.33	3.28		5.38	31.54
5.	Im		8.54	17.52	5.56			32.42
	Ee		7.01	16.05	5.47		4.41	33.34
	Es		6.47	15.46	6.02		3.36	32.11
6.	Im		16.16	11.53	4.27			32.36
	Ee		14.08	11.64	4.28		2.16	32.56
	Es		14.28	11.37	4.25		2.19	32.49

to the strain on the student when he is reduced to aural signals only, all the more so as this is not even continually the case when he receives instruction in his own mother-tongue (1964, p. 105). On the other hand communication between individuals involves both the listening and the speaking skill, and the faculty of listening and understanding has to be trained as well as other skills.

The audio-lingual method was also followed in that there was always an oral introduction of the matter to be learnt. That does not mean that the individual lessons always started with an oral section. They could also start with a written follow-up of what had been discussed during the preceding lesson.

The Explanations.

The total time for the grammatical explanations during the six experimental lessons was in the Ee group 26 minutes and 10 seconds and in the Es group 24 minutes and 31 seconds. Thus in the Ee method the explanations were more extended in time in spite of the fact that Es includes more features than Ee (see table 2). One reason for the longer time in the Ee group could be that explanations in a foreign language must be given more slowly and distinctly than explanations in Swedish.

As mentioned earlier the explanations were placed where the student was considered to profit the most by them. Their length was also to depend upon the problem in question. It was believed that in adhering to what was deemed to be the most useful in a particular situation, optimal explanations would be arrived at. Naturally, under these conditions, the explanations were differently spaced in the separate lessons, but within the same lesson in the explicit variants they occurred with few exceptions in the same place. The frequency of the explanations also varied from twice to five times within the separate lessons, and the length showed variations from 15 seconds to two minutes and 39 seconds (see table 3).

The content of the explanations focussed on how passive sentences are formed. The problem on when the passive voice is used in preference to the active is only mentioned in passing on a few occasions.

The grammatical content of the explanations will be clear from a study of the survey on page 39 and 40.

Table 3 : Occurence and Length of the Explanations in the Ee and Es Strategies.

Lesson 1	Ee:	0.20	0.53	1.25		
	Es:	0.21	0.42	0.54	1.16	
Lesson 2	Ee:	1.05	1.50	0.23	1.00	
	Es:	0.50	1.45	0.30	1.00	
Lesson 3	Ee:	1.28	2.22	0.24	0.27	1.36
	Es:	1.18	2.05	0.21	0.15	1.41
Lesson 4	Ee:	2.34	2.16	1.10		
	Es:	2.38	1.59	1.01		
Lesson 5	Ee:	2.39	0.16	1.46		
	Es:	2.06	0.18	1.12		
Lesson 6	Ee:	1.15	1.01			
	Es:	1.15	1.04			

The nature of the explanations varied very much as some could consist of a simple observation on a crucial element in a sentence while others were made up of a fairly lengthy theoretical analysis of how relations between active and passive sentences functioned. The explanations could also consist of a follow-up of a theoretical discussion in a written exercise.

Instructional Content per Lesson.

The lesson materials in the present study is to a certain extent identical with the materials of GUME 3. However, as a higher form was chosen for the GUME 5 experiment, the grammatical content as well as the lesson materials was enlarged. Below follows a survey of the content of each lesson. It should be noticed, though, that as all lessons include transformations from passive to active and vice versa as well as elicitation of passive sentences by means of various stimuli, this has not been mentioned under each heading.

Survey of the Lesson Content.

Lesson 1: *THE PRESENT TENSE*

A. Revision of passive constructions dealt with earlier.
B. Number.
C. Contrasting passive and active sentences.
D. Introduction of transformations of passive sentences with a logical subject.

Grammatical Content:

a. Comments on number.
b. Remarks on why sentences are called "active" or "passive".
c. Observations on the fact that passive sentences need not always state the logical subject of the sentence.

Lesson 2: *THE PAST TENSE*

A. The present and the past tense contrasted.
B. Number.
C. Introduction of transformations of passive sentences without a logical subject.

Grammatical Content:

a. Schematic illustrations of transformations (by means of arrows, etc.).
Comments on the fact that the logical subject of the active sentence is still the logical subject of the corresponding passive sentence.
Analysis of the verbal part in the active and the corresponding passive sentence.
b. Discussion of number.
c. Observations on passive sentences without a logical subject.

Lesson 3: *THE PERFECT TENSE*

A. The past and perfect tenses contrasted.
B. Number.

Grammatical Content:

a. Schematic illustrations of transformations. (Use of passive sentences with and without a logical subject. Auxiliaries and the principle parts of the verb discussed).
b. Comments on number.

Lesson 4: *THE PLUPERFECT*

A. The perfect and the pluperfect tenses contrasted.
B. Number.
C. Functions of the infinitive and the past participle demonstrated.

Grammatical Content:

a. Schematic illustration of transformations.
   Observations on the verbal part of passive sentences.
b. Remarks on the principal parts of the verb.
c. Observations on number.

Lesson 5: *THE FUTURE TENSE*

A. Revision of total tenses practised.

The Grammatical Content:

a. Analysis of the verbal part of active and passive sentences.
b. Observations on transformations.

Lesson 6: *THE MODALS CAN AND MUST + THE PASSIVE INFINITIVE*

A. Revision of the future tense.

Grammatical Content:

a. Observations on transformations.
b. Discussion of the verbal part of passive sentences.

THE GUME 5 PROJECT
A DESCRIPTION OF THE LAY-OUT

Objectives.

Although the research strategy was modified in some respects as a consequence of earlier results (see the next section), the main objectives remained almost the same as those for GUME 3.

1. to investigate what effects theoretical explanations in juxtaposition to pure structure drills may have on learning as compared to drills without explanations
2. to compare learning effects when
   a) explanations are given in the target language (English)
   b) explanations are offered in the source language (Swedish) and comparisons made with it
3. further production of various sorts of achievement tests in English
4. continued production of instructional materials.

In the main the present report will deal with points 1 and 2.

The Present Investigation in Relation to Earlier Results.

As was mentioned earlier, the present study was modelled on the GUME 3 study. In the latter experiment the lessons were strictly divided into three parts, namely speaking, writing, and reading. The time set apart for grammatical explanations in the Ee and Es groups was then about one third of the 30 minutes of instruction. Thus three minutes were taken from each of the three modules and given to grammatical discussion and analysis. The choice of nine minutes each lesson for explanations was made for experimental purposes and not on grounds of what was considered customary during the English instruction in grade 7. The follow-up project, GUME 5, meant several changes. First, the experiment was moved up to the grade above, grade 8, and secondly, no strict division of the time factor was undertaken in advance. Consequently, the time devoted to explanations was to depend on what was deemed appropriate in a particular situation and not on principles decided on in advance.

It was postulated that the choice of grade 8 would mean two things. There are indications that in the teaching of grammar to mature students and probably in an increasing degree to anyone over twelve, explanations are beneficial to the learner (Campbell, 1970, p. 45, and Rivers, 1968, p. 81). Therefore it might be expected that the pupils of grade 8, being one year older than those of grade 7, would profit comparatively more by grammatical explanation. The second implication which the choice of grade 8 entailed was that the teaching materials, being the same in sk and ak, and striking by necessity an average note between the two courses, would give rise to some dissatisfaction in both groups. The difference between the courses is not static through grades 7-8 as, for instance, more pupils leave ak for sk than vice versa. (In 1968 the sk group consisted of 65 % of the total population in grade 7, whereas in 1969 the same batch had grown to 71.8 % in grade 8; Statistiska Meddelanden U 1969:5 and U 1970:5). Furthermore, the divergence is augmented due to the well-known fact that talented pupils make comparatively more progress than the less talented ones (Anastasi, 1958, p. 211).

Another change in the design of the GUME 3 experiment was that the earphones used by the pupils during the instruction were dropped in the present study. All instructions were given from a tape-recorder. The decision to have "canned" lessons remained, though. The role of the teacher, which during the previous study was reduced to that of a passive maintainer of discipline, was a little more active in GUME 5. He was still not to answer questions on problems connected with the instruction, but he was to lead chorus reading with gestures, help the students find the right place in their papers, and give the right answers in the so-called free conversation exercises.

An instructional period of six lessons is a very short time to base any conclusions about progress on. Nevertheless, it was deemed unrealistic to try to obtain more lessons than six in grade 8 for the experiment. In this particular grade the pupils leave the school premises for three weeks at intervals for practical vocational guidance. As only half the class leave at a time for offices, workshops, etc., the consequence is that for about six weeks of the school term no ordinary teaching, i.e. with the whole class assembled, can take place. Besides, grade 8 has only three lessons of English a week as compared to the four lessons a week in grade 7.

Time Schedule.

GUME 5 had been scheduled to start in March 1970 (week 11). According to the time table made up for the project (see figure), the test of scholastic aptitude (DBA), PACT, and the Pre-test were to be administered before the Easter holiday. After Easter the series of six lessons was to start and as far as possible be completed during weeks 14 and 15. According to plans, the Post-test was to be administered in week 16, and an Attitude test after the experiment was completely finished.

Figure 2 : Chronology of GUME 5.

Week 11	DBA	PACT	
Week 12	Pre-test	Pre-test	
Week 14	Lesson 1	Lesson 2	Lesson 3
Week 15	Lesson 4	Lesson 5	Lesson 6
Week 16	Post-test	Post-test	Attitude test

The month of April had been chosen for the instructional part of the project because no official holiday occurred in that month, which meant that the ordinary number of lessons was not diminished for that reason. It had been anticipated that delays in the plans might occur owing to medical examinations of the pupils, football championships between and within the schools, sports days, lectures to all the school, and the like, and our apprehensions in these respects were confirmed. Very few classes had brought the project to an end within April. During the month of May only 16 days could be reckoned with as school days, and furthermore the teachers were then obliged to administer the standardized national tests for grade 8 in English, something

which required three lessons. In a few cases extra periods for the experiment were obtained and about the middle of May all test papers had been collected from the schools.

The Teaching Methods.

The experimental treatments (independent variables) used in the study are nominally the same as those used earlier, namely the Implicit and the two Explicit methods, abbreviated

Im
Ee
Es

However, since there are certain discrepancies between these methods and those used in GUME 1-3, and since interpretation of the results is dependent on a clear picture of what "happened in the classrooms", we gave a rather detailed description of the methods in the previous chapters.

The Experimental Population.

In each of the first three GUME studies 18 classes had been used in the field work. Six of these classes were from ak and twelve from sk. The distribution of the pupil sample with one third from the ak classes and two thirds from the sk classes had been decided on because it was at that time the proportion when pupils in grade 6 made their course choice for grade 7. During the experiments it was found that as the ak classes normally had fewer pupils and higher figures for absence, the desired proportions between the streams were upset. For the present investigation it was therefore decided that 12 classes from ak and 12 classes from sk should be included.

To enlist classes for GUME 5 proved to be much more difficult than it had been in connection with the earlier studies. Class 8's which had taken part in the experiment in grade 7 were excluded as a matter of course. To secure teachers of grade 8 for the project a presentation to be distributed among them was sent out to a great number of head-masters of comprehensive schools in Göteborg. In their answers some teachers stated downright that their ak classes were unsuitable to take part in any experiment for disciplinary reasons. Teachers of sk classes sometimes gave as a reason for their refusal to participate the low number of lessons in English in grade 8 which made it impossible

for them to desist from leading the instruction themselves. The search for suitable classes which was started in November was thus not terminated until the end of February. It then included 13 schools, of which nine were in the east, northeast and southwest parts of Göteborg and two in Mölndal, a town close by Göteborg. Two schools in municipalities in the county of Bohuslän also had to be included since the pursuit of ak classes in Göteborg had only secured ten classes. A list of participating school classes will be found in Appendix D .

It may be argued that in an experiment like the present one it would have been better to concentrate on one of the courses - sk or ak - trying to optimize the teaching materials for that course rather than making something intermediate and non-optimal for both. This way of reasoning is particularly relevant when the main purpose is to produce teaching materials; it should be remembered, however, that in the present investigation the main objective is to investigate whether explanations facilitate the learning of certain grammatical items, which might be done with less than perfect materials. Furthermore, in the light of Lgr 69, where it is stated in so many words that the goals for sk and ak in English are the same (p. 145), it becomes of great interest to investigate if one and the same teaching materials can function in both courses. However, from our viewpoint, the most negative consequence of including both courses is that they become limited in size (the total number of classes was what our resources permitted).

Assignment to Treatments.

Within each course the 12 school classes were randomly assigned to teaching methods. However, one restriction was applied to this procedure: no two classes from the same school were allowed to get the same treatment. Incidentally, the randomization procedure was undertaken on March 10th, 1970, shortly before the beginning of the project and after all materials were written and the teachers informed about the project.

Drop-out Rate.

The twenty-four classes participating in the experiment contained 519 pupils in all. However, ninety-four of these missed either the Pre-test or the Post-test, or both, and were therefore excluded from

the study, leaving a rest of 425 pupils. As a matter of fact, for the ninety-four pupils just mentioned no cards were punched. Although information was available on these pupils in a number of variables, it was decided on that only pupils who could be utilized in the main investigation (treatment comparisons) should be included in the data processing. Of the 425 thus left for the experiment, thirty-eight were absent from more than one lesson and were therefore eliminated from the computations, which leaves 387 pupils. In the present report the pupils who were eliminated because of a too high rate of absence will hereafter be referred to as the drop-outs. The following table gives the details with respect to the two courses.

Table 4 : Number of Pupils Participating in the Study.

	Total N in experimental classes		No Pre- or Post-test result	"Rest 1"		Absent >1 lesson	"Rest 2"	
Sk	297	519	40	257	425	22	235	387
Ak	222		54	168		16	152	

The Criterion Test

The progress criterion consisted in GUME 5 of a test especially constructed for the project and tried out before the start of the experiment. The test consisted of 94 items and was made up of six sub-tests and was to be distributed on two different occasions. The first part, consisting of three sub-tests, took about 24 minutes and the second part, equally of three sections, lasted for 30 minutes. The same test was administered before and after the experiment.

It is a desirable quality of a test battery that the scoring can be done in an objective way. If, however, that means that productive skills are excluded from evaluation, it is questionable whether the price for objectivity is not too high. It had been decided from the start of the project that the progress test should contain tests of a receptive as well as of a productive nature. As the test, for experimental reasons, was somewhat overloaded, it could well afford having one or two parts less objectively scored than the rest. It is also evident that objective tests are not the most suitable in all areas of linguistic activity (Lado, 1965, p. 35). The decision to have some tests fairly subjectively scored was accompanied by a resolution to have re-marking performed of a randomized sample of the tests in question to establish their reliability.

The design of a test battery is necessarily decided by the content matter which the student is supposed to have learnt. Thus the passive voice in the different tenses, which made up the lesson materials, was bound to be included. Furthermore, as the teaching had stressed the interrelationship between active and passive sentences, and the formation of the verbal part of passive sentences, these were also essential ingredients of the progress test. Designing the test also meant a decision about which of the four fundamental skills, listening and understanding, reading and understanding, speaking, and writing were to be chosen to test the above elements in each individual part of the progress test.

To reach as great uniformity as possible in the different classrooms the testing procedure was regulated from a tape which gave all instructions and information to the pupils in Swedish. The tests had been graded as far as difficulty was concerned, so that the easiest tests were administered at the first testing occasion and the more advanced ones on the second occasion.

Description of the Sub-tests.

Part 1. The first sub-test of the test battery (see Appendix A) consists of a completion test where 11 different forms of the auxiliary *be* have been removed in a running text. The student is to fill in the blanks. The density of the blanks varied which was a consequence of the crucial element chosen. There were no scoring problems, as no other word could possibly have suited the context. This type of test is considered by Lado (1961, p. 251) to be useful when measuring knowledge of grammatical structures. It is also said to test the passive ability of comprehension as well as the active skill of producing the right word in a context (Dahlgren, 1947, p. 173). As the text was about music it might be true that extralinguistic features were also tested as not all pupils take music as a subject at school, and it can be doubted whether all of them had heard of Tchaikovsky earlier. Anyhow, it was likely to strike the various methods to the same extent. The test consists of a mixture of tenses.

Part 2. Part 2 employs only the perfect tense. The students are to perform transformations on 10 sentences from the passive into the active. The form of the main verb to be employed is presented in the stimulus sentences and thus only the auxiliaries change. In this test, number is marked and no spelling mistakes are permitted.

Part 3. A listening comprehension test is the last part on the first testing occasion. The pupils listen to a text on Dr. Dolittle's adventures. The nine extracts end with a passive sentence and the pupils are to choose the corresponding active sentence among four alternatives. This sub-test is all in the past tense.

Part 4. This test, which started the second testing occasion, is identical with sub-test 4 of GUME 3. It tests reading comprehension and consists of 40 items in the future tense. The pupil is to decide if he has an active or a passive sentence in front of him and accordingly put a cross below either *will be* or *will have* in the margin.

Part 5. The test consists of 16 English sentences which the student is to complete. This type of test is called partial sentences items by Valette (1967, p. 141). The difference between completion tests

and partial sentences tests is that in the former case the student does not receive any cue as to what his answer should be. As test 5 is to measure the students´ ability to form the verbal part of the passive sentence, the verb to be employed is given in the infinitive.

The right answers to the first two questions are dictated to the pupils, and they hear the verbal part three times. In consequence these two sentences were not scored. This dictation involves a listening discrimination element in an otherwise written test, and the mistakes made by a few pupils in the test papers (they wrote *was taking* when *was taken* was dictated to them) implies that impaired hearing, outside noises, position in the classroom, or tensions could have influenced a correct understanding of what their task was.

For part 5 the present and the past tenses were chosen. The tense had to be correct for the pupil to receive a full point, but number was not scored this time and no "unimportant" (*bee* instead of *be*, *whas* instead of *was*, etc.) spelling mistakes either.

Part 6. The last test is an entirely written test. It consists of 10 items which are made up of the same drill pattern as the pupils practised during the lessons, that is transformations between active sentences and the corresponding passive sentence without a logical subject. The only difference is that in the instructional situation the transformation was always passive/active while section 6 tests active/passive. The tense of this particular test is the present in all the sentences but the last which employs the past tense.

Of the six part tests four were productive and two receptive. No test excluded reading comprehension, and no test was of an entirely "pure" character so that only one element was measured in reference to one skill only. A few of the tests were of a type which the students could recognize from the experimental instruction, e.g. Nos 4, 5, and 6. The mixture of partly new and partly well-known types of tests was intentional.

Reliability. From a purely experimental point of view it would have been desirable that the pupils were completely ignorant of the grammatical structure in question. In that case a zero-point had existed from which the pupils would probably have deviated to a higher or lesser extent after the lesson series. However, from the conventional

reliability point of view the test would have been considered very unreliable on the pre-test occasion, all the pupils having the same score and the test showing no discriminative power. Apparently the conventional reliability approach is not valid in similar situations. It would rather seem that the lack of variation in scores would be an indication of reliability (as expressed in the standard error of measurement); the ignorance of the pupils has been perfectly tapped by the test, which apparently consists of very homogeneous items. Thoughts along these lines are expressed in a forthcoming book (Marton & Levin, 1971).

The test used in the present study proved not to fulfil the requirement mentioned above (the "zero point"); as a matter of fact, the grammatical point was not as unknown to the pupils as might be expected. Thus both our experimental groups scored fairly high on the pre-test. Therefore, and also because the Pre-test was used as a covariate in some of the analyses of covariance (see below, p. 61), the reliabilities (Kuder-Richardson 21) were calculated; they were as follows:

The Pre-test	sk	ak
Part 1	.53	.52
2	.78	.68
3	.69	.52
4	.74	.43
5	.70	.21
6	.72	.32
Total	.90	.58

The inter-scorer reliabilities for tests 5 and 6 (two independent scorers) were:
ak: .90
sk: .98

Despite the fact that the criterion test proved to be very easy for the sk group, it measures inter-pupil variance with great precision. In ak, however, some of the part tests have too low reliabilities even for group comparisons; however in the main investigation only the total score is used.

Validity. The validity aspect of the test will be commented on in a later chapter (p. 99 ff). Here it may be stated that the criterion test correlates (in sk) .72 with Grades English and .75 with the Standardized test in English, the corresponding figures in ak being .43 and .51 respectively.

The Pupil Attitude Test.

After the lesson series the pupils were given an attitude test. This questionnaire is given in Appendix B, but it will be discussed briefly here.

The questionnaire partly consists of multiple choice items and partly of items of the open answer type. As is obvious from Appendix B the two response types are mixed in the test.

Of the multiple choice questions, nos. 4, 5, 8, 9, 11, 13, 14, and 15 are added together to a total. This total measure is supposed to reflect the pupils' general attitude to the project, although the various items concentrate on different aspects of it. The first four items are 5-choice, the last four 4-choice; thus the maximum score is 36.00. The theoretical mean is 22.00, indicating a neutral attitude to the experiment in general. The eight items with fixed response alternatives focus on the following things:

4: the pupil's own idea of whether he made more or less progress in the experiment than during ordinary lessons
5: if the pupil thought that the experimental lessons were more or less fun than ordinary lessons
8: if time passed faster or more slowly than during ordinary teaching
9: if the student was more tired or less tired than after ordinary lessons
11: the sound quality of the tapes
13: the oral exercises
14: the written exercises
15: the reading texts.

The only remaining multiple choice item in the questionnaire, No. 12, referred to the explanations given in the Ee and Es methods. Since the Im classes were supposed not to answer this question, the item was excluded from the total Attitude score. However, it will be treated separately in the results section.

The open answer items asked the pupils to comment on: what was good about the GUME experiment (2), what was not so good about it (3), what was fun in the lessons (6), what was boring (7), what had made the pupils feel tired - if tired they were (10), and finally, item 16 asked them to make whatever additional comments they wanted.

The pupils were also asked to indicate their interest in the various school subjects (the first page of the questionnaire). The intention behind this item was mainly to find out how the pupils ranked English in relation to other subjects.

The Teacher Attitude Test.

The questionnaire consisted of two parts. In the first part the participating teachers were asked questions on how they usually teach English themselves, which method they use (as compared to those used in the project), how they treat grammatical points, how much they speak English, etc.

In the first part of the questionnaire some background information about the teachers was also collected.

The second part of the questionnaire required the teachers to comment on various aspects of the lesson series: the grammatical explanations (in the Explicit groups), the oral exercises, the written exercises, the reading passages, the tempo of the lessons, the sound quality of the tapes, the reactions on the part of the pupils, etc. Further particulars about the questionnaire will be given in connection with the presentation of results (see p. 107 below).

The Standardized Test in English.

All Swedish pupils in grade 8 are given standardized tests in Swedish, English and Mathematics, prepared by the National Board of Education. The English test has been used for many years and is somewhat out of step with recent developments in language instruction. Incidentally, members of the GUME project have been contracted for research on and development of new tests in this particular field.

The tests are put at the teacher's disposal to help him arrive at the greatest possible uniformity when assessing the standard of the individual class in relation to the national norm; thus the tests are not primarily aimed at indicating the individual student's standing.

The tests were administered in May, towards the end or after the present investigation. Below follows a brief description of the various parts.

Sk (the advanced course):

EL - (Läsförståelse), Reading Comprehension (35 minutes)
EM - (Meningsprov), Vocabulary Test (28 minutes)
EA - (Avlyssningsprov), Listening Comprehension (20 minutes).

EL is a reading comprehension test of the conventional type. The pupils read minor passages (somewhat longer in sk than in ak), and get three to six multiple choice questions (5-choice) on each passage. There were 31 questions in all.

EM is a kind of active vocabulary test. The pupil gets a stimulus sentence, as, for instance, "Tom is a boy, Alice is a x x x x", and is required to mark one of the following letters: G A Z O F, one of which is the first letter of the correct word. This part consists of 28 items.

EA is a listening comprehension test. The pupils listen to a spoken passage and mark on a separate answer sheet four multiple choice questions (5-choice) on each passage. The choices consist of written alternatives. There are five such passages and in all 20 questions.

Thus the complete test contains 79 items altogether.

In the easier course, ak, the test variables have the same denominations as in sk and the test types are the same. However, the content is different and in line with standards in the ak population. The particulars of the ak variety of the test are given below:

EL:	35 minutes	20 items
EM:	28 minutes	20 items
EA:	18 minutes	28 items
		Total: 68 items.

The three tests are given in three different class periods, normally on three different days.

The reliability of the total test is: sk = .90; ak = .86.

PACT.

The original test, called Pictorial Auditory Comprehension Test, was developed by John B. Carroll and one of his assistants, Wai-Ching Ho.

It is a listening comprehension test intended to measure foreigners' comprehension of spoken English. In GUME 1-3 mimeographed copies of the original version were used by kind permission of Dr. Carroll. In GUME 4 (grade 6) as well as in the present study, entirely new versions were worked out, although with the original testing technique preserved. As a matter of fact, the version used in GUME 5 was made before the one used in GUME 4; the latter one has been accounted for in the previous report (Lindblad & Levin, 1970, p. 52). The version used in this study should thus not be compared with the one in GUME 4; in fact the GUME 5 version proved to be much easier. Since the development of a listening comprehension test has been a specific objective within the GUME investigations, a few words of comment may not be out of place:

The means for PACT, i.e. the easy version, in GUME 5 (grade 8) are: sk: 50.19; ak: 43.19. These values should not be compared with the mean obtained in GUME 4 (grade 6), which was 34.29. However, the latter and more difficult version has also been tried out in grade 8, the sk as well as the ak course, although on small samples so far (N = 50 and N = 35 respectively). The means there were: sk: 39.88; ak: 32.77. The results indicate that, on the average, the entire population in grade 6 scores higher on this particular test than do the two years older ak pupils.

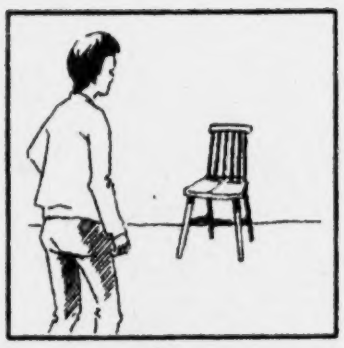
In this test the pupils listen to a taped conversation or description of an object or event, etc., and then mark which of four alternatives (in the form of pictures) that corresponds to what is said on the tape. The test consists of 55 items and takes 25 minutes to administer. The reliability (K-R 21) of the test is: sk = .55; ak = .92. As it appears, the test is not very reliable for the advanced course; it proved to be too easy.

Although auditory tests have been available in the Swedish schools, none have been uncontaminated as far as reading ability is concerned (the options on the answer sheet have mostly consisted of w r i t t e n alternatives). PACT seemed to be promising in this respect and was therefore further explored in the project. On the next page the testing technique is illustrated by an example.

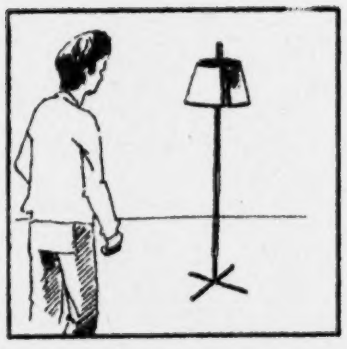
gumeprojektet

1

a



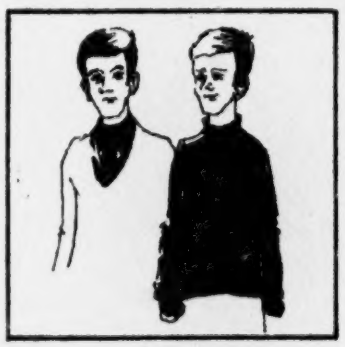
b



c

d

2



a

b

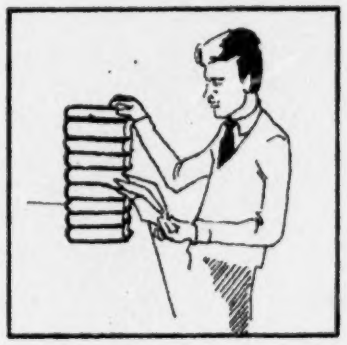
c

d

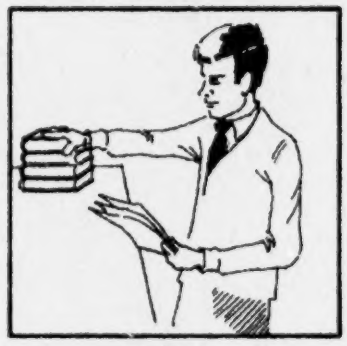
3

a

b

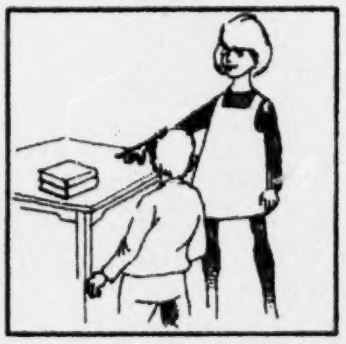


c



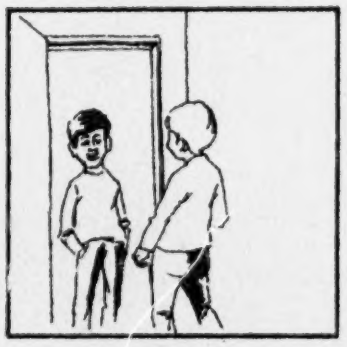
d

4

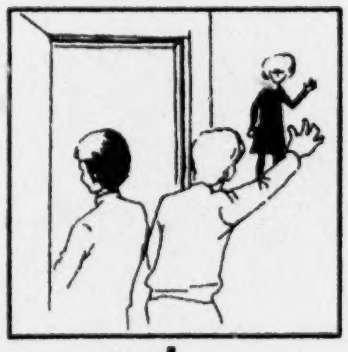


a

b



c



d

PACT.

The four first items of the test are presented above.
As a typical example the auditory stimulus of item No. 4 is given (the following is heard from the tape):

"He´ll come when he´s finished his homework".

The pupils mark their answers on a separate sheet.
(It is B which is correct, of course).

General Scholastic Aptitude (DBA).

In the present investigation the same test as was used in GUME 1-4 was administered, namely the verbal, inductive, and spatial parts of the so-called DBA test (DBA = Differentiell BegåvnirgsAnalys, i.e. differential intelligence analysis) developed by Professor Kjell Härnqvist of the University of Gothenburg. In actual practice, the test is used mainly as an aid in vocational guidance. The three subtests, taken together, are considered to be a reliable measure of general ability or scholastic aptitude (see further in Härnqvist, Manual till DBA). The sum of the pupils' three stanine scores were transformed to T-scores with a theoretical mean of 50 and a standard deviation of 10.

The tests were given two weeks before the experiment proper started (see figure 2, p. 43). The verbal and inductive parts were given on one occasion, the spatial part on a second (when also PACT was administered).

The reliability for the total test was in this study: sk: .65; ak: .59. The relatively low reliabilities are of course explained by the fact that they refer to two groups, both homogeneous in comparison with the total population.

Other Measures.

Social class. Information about the parents' occupation was collected at the headmasters' offices. The intention was partly to check the social background of the different treatment groups and partly to investigate the correlation between this variable and others used in the study. The criterion for assigning a pupil to a particular social class was a hierarchical description of professions and occupations from 1958 (1958 års valstatistik), which is to some extent arbitrary and even inconsistent, but it is the only source available at the moment. Social class 1 corresponds roughly to English "upper middle class", and class 3 to "working class"; the much disputed division is based on income only. A zero was used as a code for cases where the mother (without any mention of profession) was given as the guardian in order to make further analyses of this group possible.

Grades. Grades in English, Swedish, and Mathematics were collected. The grades had been given at the end of the term preceding the

experiment, i.e. when the pupils had finished the first term of grade 8. It should be noticed that the grades had not been corrected or adjusted according to any standardized achievement test, simply because no such test had been given (the last time this happened was towards the end of grade 6, that is when the total group had not yet been streamed into two courses). Thus the grades in this investigation reflect a certain amount of subjectivity. They are expressed on a 5-point scale (theoretical mean 3 and standard deviation 1). The three grades were added together whereby a scale with a standard deviation of 3 was obtained; these values were in turn multiplied by 3 in order to make them comparable to the DBA scores, which are expressed on a scale with about the same standard deviation. The purpose of the procedure was to give Grades and DBA equal weight in the statistical analyses. The addition of the three grade scores is somewhat dubious since the pupils do not belong to the same reference group in the three subjects. However, this will be commented on further when the populations are described with respect to grades (see p. 69 below).

Grades German and French.

From grade 7 and onwards the pupils have the possibility of studying a second foreign language, German or French. Practically all pupils belonging to the sk group in English choose one of these languages. Among the ak pupils forty-four took German and one pupil French. In each language there are two courses, one advanced and one easier.

The intention behind collecting the German and French grades is to compare, in cases where this is possible, the correlations between the various language tests on the one hand and English, German and French grades on the other.

Comments on an Oral Test.

An oral test was also added to the test battery. It was administered after the Post-test had taken place and included 12 classes of the total pupil sample. Six of these classes were from ak and six from sk.

Very few schools in Göteborg have as yet a language laboratory, which meant that the test had to be administered individually in the

schools by an assistant. The test, which had been modelled upon sub-test 5 of the Criterion test, takes five minutes for each pupil.

A forthcoming report will give a thorough description of the oral test. It will also compare in detail the results of the oral test and the corresponding written test, sub-test 5. An error analysis is in progress and will later on involve an investigation into the degree of acceptability of some of the responses given by the pupils.

DATA TREATMENT

The Statistical Program.

All data were processed at Göteborgs Datacentral för Forskning och Högre Utbildning by computer IBM 360/65. Statistical programs included in the ISR (Institute for Social Research, University of Michigan) and BMD (Bio-Medical Computer Programs, UCLA) series were used. The following measures or analyses were obtained:

a) Means, standard deviations and frequency distributions for all variables. These data were obtained for the total population, for the sk and ak groups separately, for boys and girls separately, for the treatments (Im/Ee/Es) separately and for each participating school class.

b) Correlations between all variables for the sk and ak groups separately.

c) Analyses of variance (one-way) on a number of independent variables in order to investigate comparability between the treatment groups (three cells).

d) Analyses of variance (two-way) with the total experimental group (sk + ak) divided into three levels of scholastic ability (nine cells).

e) Analyses of covariance with different covariates and dependent variables.

The purposes of the various analyses will be given below. A pupil not attending 5 or 6 lessons was eliminated from the data processing. In a field study of the present kind it is necessary to accept a certain amount of absence, otherwise there is a risk that the sample will be limited severely. As in the case of the earlier GUME studies where the treatment contained six lessons, the line was drawn at one lesson as an acceptable amount of absence. The pupils who did not take the Pre-test and the Post-test were also eliminated from all computations, even if they had taken part in the whole lesson series. Within the two experimental populations (sk and ak) the N's vary somewhat from variable to variable due to stray absences.

Experimental Design.

The design corresponds to Campbell and Stanley's "design 10", The Nonequivalent Control Group Design (Gage, 1963, p. 217). For administrative reasons intact school classes had to be used in the experiment. It has thus not been possible to assign pupils randomly to teaching strategies (treatments). In the absence of experimental control of background variables, statistical control by analysis of covariance has been resorted to when investigating the main effects.

The unit of analysis used in the experiment is the individual score. Since it might be argued (Wiley, 1969, p. 213) that the school class mean should be the proper unit of analysis, an investigation of the main effects has also been made in accordance with this view. Of course, with such a limited number of school classes as are used in the present investigation, the loss of degress of freedom is great when one moves from the individual to the school class level.

Investigation of Main Effects.

The main purpose of the experiment is to investigate which of three teaching methods produces the best learning results. The measure of progress that was used through the computer analyses was the difference in raw scores between the Post-test and the Pre-test. In addition, two other measures of progress were used though in those cases the computations were made by hand. The particular measures will be presented below.

When the three teaching strategies were compared with respect to Progress, the following covariates were used in the analyses of covariance: DBA, PACT, the Pre-test, and the Standardized Test in English. The four measures were used separately in four different analyses; in a fifth analysis they were weighted together to a composite measure. Treatment effects were also compared with respect to Post-test scores; in this case the Pre-test served as the covariate. The analyses of covariance may be summarized thus (the analyses were identical in sk and ak):

Survey of Analyses of Covariance Performed in sk and ak.

Dependent variable	Covariate
PROGRESS	DBA
-"-	PACT
-"-	The Pre-test
-"-	The Standardized Test in English
-"-	The above four weighted together
THE POST-TEST	The Pre-test

Investigation of Interaction Effects.

Our intention here is to divide the total experimental population (sk and ak) into three groups according to DBA scores. This procedure is of course somewhat dubious since the two groups consist of pupils who have, for three terms of English, been taught separate courses. They are thus very distinct groups and should not be treated together in the statistical analyses. However, the procedure may be accepted as a tentative investigation of whether any interaction exists between scholastic aptitude and teaching method in the larger population. The DBA scores (ranges) for the upper, middle, and lower third turned out to be: 26-47, 48-56, 57-74. The data were organized in a 3 x 3 table, thus:

	Im	Ee	Es
Upper			
Middle			
Lower			

Retention.

According to the original research plan our Criterion test should be administered a third time, when the pupils were just starting grade 9, in order to measure retention or, rather, differential retention between the three methods. (In GUME 1-3 the retention tests were given one

month after the experiment). However, for the results to be interpretable it would have been necessary to control the teachers for an unduly long period of time, preventing them from teaching the structures dealt with in the project. Since it was considered unrealistic to control the teaching process in this way, the retention test was dropped.

Various Measures of Progress.

As has been mentioned earlier, the pupils' progress during the experiment was measured by the difference in *raw scores* between the Post-test and the Pre-test. However, it may be argued that a measure of progress must somehow take account of the pupils' standing on the Pre-test. If, for instance, a pupil scores very high on the Pre-test, there is not so much room for progress because of ceiling effects. The following index takes care of this, giving more weight to progress "at the upper end of the scale":

$$\frac{\text{Actual improvement} \times 100}{\text{Possible improvement}} = \%$$

An example: Pupil A has 70 points on the Pre-test and 90 on the Post-test, pupil B has 50 on the Pre-test and 70 on the Post-test. The improvement of both these pupils is thus 20 points and according to this measure they have made the same progress. The Criterion test used in the present study has a maximum score of 94. Possible improvements for the two subjects are 24 and 44 points respectively. Their scores as computed by the above formula then become 83.3 (%) and 45.4 (%) respectively; thus pupil A has made greater progress according to this measure.

On the other hand it may be argued that increments among inferior pupils are of greater consequence than equally great improvements (in raw scores) among superior pupils. However debatable this way of reasoning may be, the following index of progress gives higher credit to improvements "at the lower end of the scale":

$$\frac{\text{Progress} \times 100}{\text{Pre-test}} = \%$$

Both these measures have been calculated with the school class mean as the unit of analysis.

STATISTICAL DESCRIPTION OF THE EXPERIMENTAL POPULATION

Attendance.

One criterion for including a pupil in the data processing was that he had been present during at least five out of the six lessons. In the table below the experimental population is described with respect to attendance during the series of lessons.

Table 5: Attendance of the Experimental Population (sk + ak) during the Series of Lessons.

	Number of lessons attended		
	6	5	
Sk	194	41	235
Ak	112	40	152
Total	306	81	387

As the table indicates, the attendance is proportionately greater in sk than in ak, which is according to expectations. For the purposes of the experiment, the pupils who were absent one lesson were considered comparable to those who had 100 % attendance (= 6 lessons). As a partial check on this proposition, absence was included as a variable in the calculation of correlations. As it appeared, absence (defined as absence during 1 lesson) did not correlate with any other variable.

Boys/Girls Ratios.

According to official statistics for the year preceding the experiment, i.e. 1969, the relation in absolute numbers between boys and girls was not the same in sk and ak. The percentages, based on more than 92.000 pupils all over Sweden, are as shown in the following table. In the same table the corresponding figures for our sk and ak groups are given.

Table 6: Number of Boys and Girls in the Experimental Group (sk + ak) and in the Total Population.

	The total population in grade 8 during 1969				The experimental sample (sk + ak)			
	sk		ak		sk		ak	
	N	%	N	%	N	%	N	%
Boys	30,160	45.4	14,840	56.8	235	60.7	94	61.8
Girls	36,283	54.6	11,288	43.2	152	39.3	58	38.2
Total	66,443	100.0	26,128	100.0	387	100.0	152	100.0

It is apparent from the table that the experimental population contains a surplus of boys in comparison with the 1969 "norm". In the case of sk the experimental group deviates significantly (X^2 = 36.28 1 df $p < .001$), whereas in the case of ak the deviation from the expected values is not significant (X^2 = 1.61 1 df $p > .50$). It should be remembered that the supply of teachers and classes willing to participate was limited (see p. 44 above). The final sample simply consists of those classes that accepted the invitation to the experiment, no matter what boys/girls ratio they had. Considering the percentages for grade 8 on the whole, it seems as if girls tend to prefer the more advanced course; they probably choose ak only in case of relatively great lack of interest in English as a subject.

In sum: in the case of sk the experimental group contains more boys than girls, although in grade 8 in general the tendency (in sk) is the opposite. It therefore becomes important to investigate if the boys/girls ratio is equal between methods; this will be done in the section below.

Assignment to Treatments.

Since the school class was the sampling unit and since the boys/girls ratio varied from class to class, the distribution of the sexes on treatments was a matter of chance. The actual distribution for sk and ak respectively are presented in the tables below.

Table 7 : Distribution of Pupils according to Teaching Methods (sk).

	Im	Ee	Es	Total
Boys	38	47	42	127
Girls	32	45	31	108
Total	70	92	73	235

In sk the Ee method has got relatively more pupils than the two other methods; however, the difference is not significant (X^2 = 3.64 2 df $p > .10$). The boys/girls ratio is the same between methods (X^2 = .89 2 df $p > .50$).

Table 8: Distribution of Pupils on Teaching Methods (ak).

	Im	Ee	Es	Total
Boys	29	29	36	94
Girls	21	20	17	58
Total	50	49	53	152

In ak the pupils are evenly distributed on the three methods. The boys/girls relation is also the same between methods (X^2 = 1.14 2 df $p > .50$).

To sum up: the distribution of pupils among the three methods is not deviating significantly from the desired one; however, in sk there are relatively more pupils in Ee. The boys/girls ratios, discussed in the preceding section, are equal between teaching methods.

Social Class.

The table below gives the distribution of pupils according to social class for sk and ak.

Table 9 : Distribution according to Social Class (absolute numbers) sk + ak.

	Social Class				
	No inform.	0	1	2	3	
Sk	50	6	41	66	72	235
Ak	2	10	2	39	99	152
Total	52	16	43	105	171	387

The "0" group stands for cases where the mother is responsible for the care of the child. As in GUME 4, the intention was to investigate this particular group with respect to a number of variables. However, in the present study the number of cases became so limited as to make any further investigation of little interest.

In the following table the "0" group as well as the group for which no information was available have been eliminated, and the remainder, i.e. social class 1,2, and 3, have been transformed into percentages.

Table 10: Distribution according to Social Class (percentages) sk + ak.

	Social Class			
	1	2	3	Total
Sk	12.9	20.7	22.6	56.2
Ak	0.6	12.2	31.0	43.8
Total	13.5	32.9	53.6	100.0

The total experimental population seems to be biassed as regards social class. According to official statistics for Gothenburg (Andrakammarvalet i Göteborg, 1968, U 1969:2, pp. 63-69) the overall figures for social groups in Gothenburg are:

1: 8.2 % 2: 38.4 % 3: 53.4 %

The deviation from this "norm" is statistically significant. The χ^2 - value obtained is 13.74 (2 df $p < .01$). This is somewhat surprising, especially when one considers the fact that, in our sample, the number of ak pupils is relatively large. The actual relation sk/ak

in our group is 60.7 % / 39.3 %, which should be compared with 71.8 % / 28.2 %, which was the distribution of pupils in grade 8 during the preceding year (Statistiska Meddelanden U 1970:5). The reason for including as many ak as sk classes has been given earlier (see p. 44). However, it is obvious that our sk sample is a select group. The following figure shows the distribution on social class for sk and ak.

Figure 3 : Distribution of Pupils (%) according to Social Class and Course in English.

The strong relationship between social class and course choice is apparent from the figure; the fact that our sk sample is a relatively select group should not invalidate this conclusion.

In short, our experimental sample is biassed as far as social class is concerned; this is explained by the fact that the sk group contains comparatively too many pupils from social class 1 and 2. In all statistical computations, however, the two groups will be treated separately.

Scholastic Aptitude.

Three parts of the DBA-test (see above, p.56), namely the verbal, inductive, and spatial factors,will be used to describe the experimental population with respect to general intelligence or, better perhaps, scholastic aptitude (the test battery is mainly used as an aid in vocational guidance). Although the two groups, sk and ak, will be treated separately in the main investigation, the standing of the total group on the DBA-test will be given to indicate the representativity of the experimental population as a whole.

Table 11: Means and Standard Deviations on the DBA-test.
(The Parts in Stanine Points, the Total in T-scores).

	Sk + Ak			Sk			Ak		
	N	$\bar{x}$	s	N	$\bar{x}$	s	N	$\bar{x}$	s
DBA Verbal	334	5.08	1.83	214	5.87	1.63	120	3.67	1.19
DBA Inductive	334	5.14	1.99	214	5.89	1.73	120	3.81	1.73
DBA Spatial	334	5.21	2.08	214	5.59	2.08	120	4.54	1.89
DBA Total	334	50.98	10.22	214	55.24	8.80	120	43.37	7.91

The DBA-test was standardized in 1958. Since then a certain increase in raw scores has been noticed for various tests at different age levels; thus the original norms have become somewhat outdated (Härnqvist, 1969). This phenomenon and its consequences for the interpretation of the test results were discussed at some length in the GUME 4 report (Lindblad & Levin, 1970, p. 63). In the case of grade 8 the changes upwards seem to be very moderate, however, (Larsson & Sandgren, 1968, p. 88) and the results in table 11 above testify to this. As far as general scholastic aptitude is concerned, our group seems to be close to the norm, which applies to the total as well as the part tests. The sk pupils are approximately half a standard deviation above and the ak pupils approximately half a standard deviation below the theoretical mean, which is in accordance with earlier findings in the GUME project (see for instance, Levin, 1969, pp. 36-37).

In sum: with respect to scholastic aptitude our sk and ak groups are such as to warrant generalizations from our results to English sk and ak groups in general.

Grades.

The grades were given during the term which preceded the experiment, i.e. the autumn term of 1969. At that time the teachers had no standardized tests available to support their grading (tests of this kind were given after the experiment; see below),which makes the grades relatively subjective in character. In the case of Swedish the ak and sk groups take the same course and should therefore be considered as one reference group with a theoretical grade mean of 3.0, whereas in English the two groups take different courses and accordingly make up two reference groups, each with a theoretical mean of 3.0.

As it appeared, a high correlation existed between course choice (advanced/easy) in English and in Mathematics; thus the pupils in our sk group were taking the sk course in Maths in most cases. However, when this was not the case, the Maths grade was adjusted downwards by one point. Correspondingly, an ak pupil (in English) who was following the sk Maths course, got his Maths grade adjusted upwards by one point. The intention with this somewhat subjective procedure, which was applied in the limited number of cases where this was necessary, was to equalize the grades in English and Maths. The actual grade means will be given in the following table.

Table 12: Means and Standard Deviations for English, Swedish, and Mathematics.

	Sk + Ak			Sk			Ak		
	N	$\bar{x}$	s	N	$\bar{x}$	s	N	$\bar{x}$	s
Grades English	381	3.15	.97	233	3.33	1.00	148	2.86	.86
Grades Swedish	381	3.02	.95	233	3.47	.84	148	2.30	.60
Grades Maths	381	3.00	1.03	233	3.21	1.06	148	2.66	.89

With respect to Swedish, where all pupils take the same course, the total group is exactly on the theoretical mean. The difference between sk and ak is according to expectations. In the case of English and Maths it is interesting to note that sk is somewhat above and ak as much below the respective theoretical means of 3.0. These tendencies for the two courses are well-known phenomena (information from the Gothenburg Board of Education).

The grade scores were added together and multiplied by three; the idea was to increase the standard deviation so as to give the Grades Total approximately the same weight as DBA. Considering the fact that sk and ak take the same course in Swedish but different courses in English and Maths, it is somewhat illogical to add the three grade scores together; the actual effect of this procedure is to diminish the true differences between sk and ak. However, for our purposes - using the Grades total as a covariate in certain analyses - this measure is accepted. The Grades total thus obtained are (standard deviations within parentheses): sk + ak: 27.42 (7.45) sk: 29.95 (7.42) ak: 23.43 (5.50). The difference between sk and ak is highly significant ($t = 9.88$).

In sum: the total experimental population lies more or less exactly on the norm in the case of grades. Large differences between the sk and ak means exist.

The Standardized Test in English.

The various parts of the national test have been described earlier (see p. 52 above). It should be remembered that sk and ak have different tests although the designations of all part tests are the same.

Table 13: Means and Standard Deviations on the Standardized Achievement Test in English.

	Sk			Ak		
	N	$\bar{x}$	s	N	$\bar{x}$	s
EL	229	18.88	5.57	148	10.56	4.32
EM	228	19.12	5.31	148	9.68	3.03
EA	229	11.30	3.83	148	18.95	4.79
Total	228	49.41	12.81	148	39.18	10.49

The theoretical means for sk and ak are 46.5 and 45.5 respectively (no norms are available for the part tests). As it appears, sk is above and ak below the respective norms.

Sex Differences in Background Variables.

In the following table the means and standard deviations for boys and girls in sk are given.

Table 14: Means and Standard Deviations for Boys and Girls in Certain Variables (sk).

	Boys			Girls				
	N	$\bar{x}$	s	N	$\bar{x}$	s	t	sign
DBA Verbal	117	5.71	1.42	97	6.06	1.85	- 1.52	
DBA Inductive	117	6.25	1.60	97	5.46	1.79	3.29	.001
DBA Spatial	117	6.05	1.92	97	5.03	2.14	3.64	.001
DBA Total	117	56.68	7.40	97	53.51	10.01	2.58	.01
Grades English	126	3.21	1.01	107	3.47	.97	- 2.05	.05
Grades Swedish	126	3.40	.80	107	3.56	.88	- 1.45	
Grades Maths	126	3.32	1.06	107	3.09	1.05	1.72	
Grades Total	126	29.60	7.42	107	30.36	7.43	- .78	
Std Test EL	125	18.80	4.83	104	18.98	6.36	- .24	
Std Test EM	125	19.06	5.07	103	19.19	5.62	- .19	
Std Test EA	125	11.04	3.66	104	11.61	4.01	- 1.14	
Std Test Total	125	48.90	11.69	103	50.03	14.08	- .65	

In our sk group the boys are significantly above the girls in scholastic aptitude; although the difference is disconcerting as regards the representativity of the sexes in our group, it affects the three treatments similarly (cf p. 65 , table 8 above). Nevertheless, the tendency for the girls to excel in the case of verbal ability, grades (with the exception of Maths) and achievement tests of a verbal nature, are well-attested facts (see for instance, Anastasi, 1958, p. 492 ff.).

In the table below the same information is given for ak.

Table 15: Means and Standard Deviations for Boys and Girls in Certain Variables (ak).

	Boys			Girls				
	N	$\bar{x}$	s	N	$\bar{x}$	s	t	sign
DBA Verbal	72	3.78	1.21	48	3.50	1.15	1.27	
DBA Inductive	72	3.99	1.67	48	3.54	1.79	1.37	
DBA Spatial	72	4.99	1.92	48	3.88	1.65	3.31	.001
DBA Total	72	45.00	7.74	48	40.92	7.61	2.85	.01
Grades English	92	2.75	.85	56	3.05	.86	- 2.01	.05
Grades Swedish	92	2.24	.62	56	2.39	.56	- .47	
Grades Maths	92	2.70	.89	56	2.61	.91	.60	
Grades Total	92	22.99	5.40	56	24.16	5.65	- 1.25	
Std Test EL	92	10.80	4.44	56	10.16	4.14	.90	
Std Test EM	92	9.74	2.86	56	9.57	3.31	.32	
Std Test EA	92	19.96	4.52	56	17.29	4.79	3.36	.001
Std Test Total	92	40.50	10.25	56	37.02	10.60	1.96	.05

In ak the boys are significantly superior to the girls in scholastic aptitude and the standardized test in English; somewhat surprisingly, the boys are above the girls even in the verbal factor. However, the girls get higher grades, especially in the case of English.

A summing of what is found in tables 14 and 15 demonstrates that the experimental population is normal as regards general scholastic aptitude (DBA) and grades. It is, however, biassed as far as social background is concerned, which is explained by the fact that sk is a relatively select group. On the national test in English the sk group is above the norm for sk, and the ak group below the norm for ak. On all measures the sk group is significantly above ak, which is according to expectations; in both groups the boys are superior to the girls in the case of DBA, although the girls excel in grades. All in all, the two experimental groups are not ideally representative samples of the respective populations. However, the biasses in the experimental group seem to have been equally distributed among the teaching methods, and should thus not affect them differently.

In all forthcoming analyses the two groups, sk and ak, will be treated separately.

Characteristics of the Treatment Groups.

In order to control the standing of the three treatment groups in the variables which are used later as covariates comparisons between the groups were made by analysis of variance. The results of the comparisons are given in the tables below.

Table 16: Analyses of Variance (One-Way) of Treatment Group Means in Certain Variables; sk.

	Means				Sum of squares			
	Im	Ee	Es	F-ratio	Between	Within	df	sign
Pre-test	60.69	60.22	56.19	2.084	908	50546	2/232	-
DBA	57.28	54.07	54.73	2.635	402	16103	2/211	-
PACT	49.99	50.90	49.51	4.315	81	2091	2/224	.05

In sk there is no significant difference between the methods in the case of the Pre-test and DBA total, although the tendency is for the Implicit method to be ahead of the others. The Ee group scores the highest figure on PACT, the F-ratio is significant despite the small differences in absolute numbers between the method means.

Table 17: Analyses of Variance (One-Way) of Treatment Group Means in Certain Variables; ak.

	Means				Sum of squares			
	Im	Ee	Es	F-ratio	Between	Within	df	sign
Pre-test	33.12	30.80	30.68	1.930	191	7381	2/149	-
DBA	42.91	41.65	45.76	2.781	338	7115	2/117	-
PACT	43.65	40.54	45.34	6.583	549	5674	2/136	.01

In ak the Es method is above the two others in the case of DBA and PACT; in the latter case the F-ratio is significant.

In sum: no clear pattern is discernible in the analyses. The most noteworthy fact is that, in ak, there is a tendency for the Es group to be ahead of the two others.

MAIN RESULTS

Overall Progress during the Experiment.

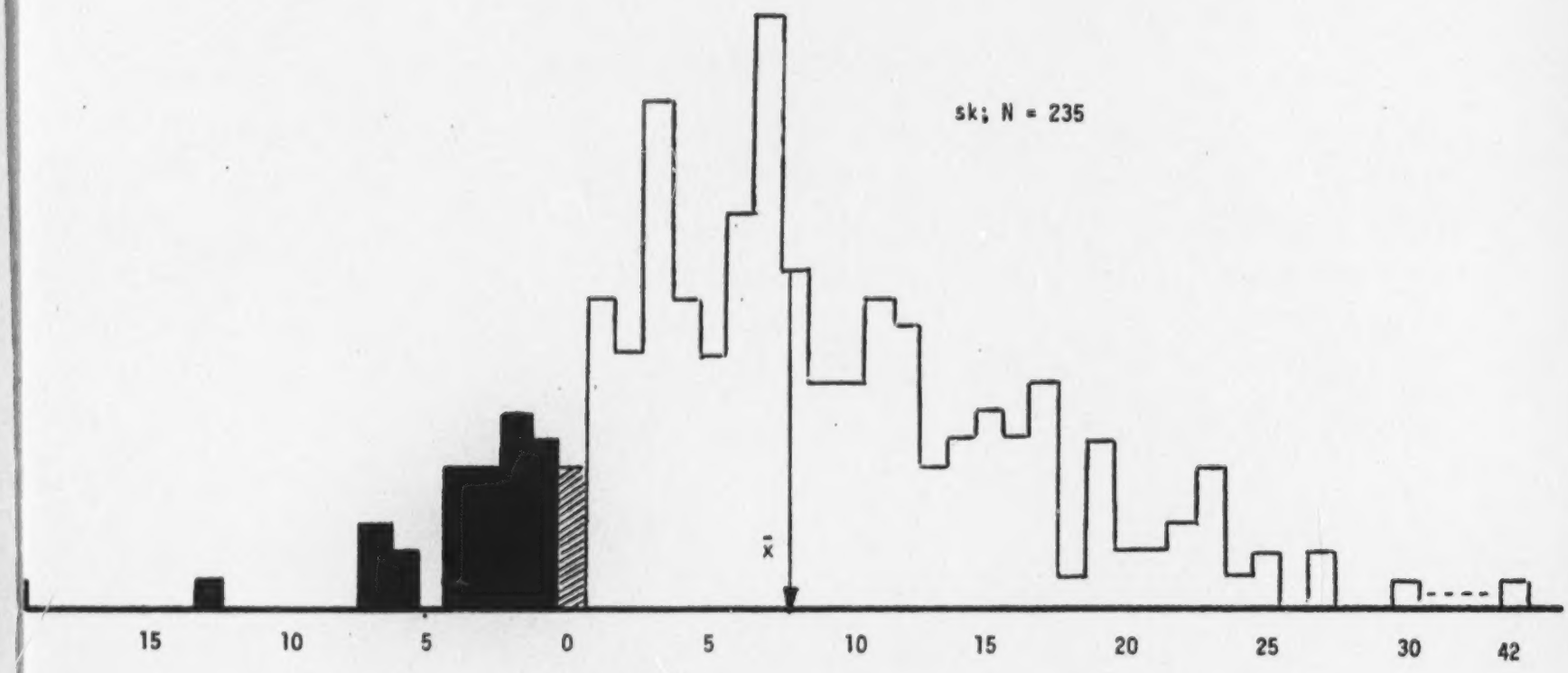
In an investigation of the present kind the main interest is of course tied to the question of "differential progress", i.e. the question whether any of the teaching methods under comparison proved to give better learning effects. However, a necessary prerequisite for investigating differences between learning outcomes is that the treatments have had measurable effects on the pupils. In other words, did the pupils learn anything, irrespective of method? The following two tables give the facts.

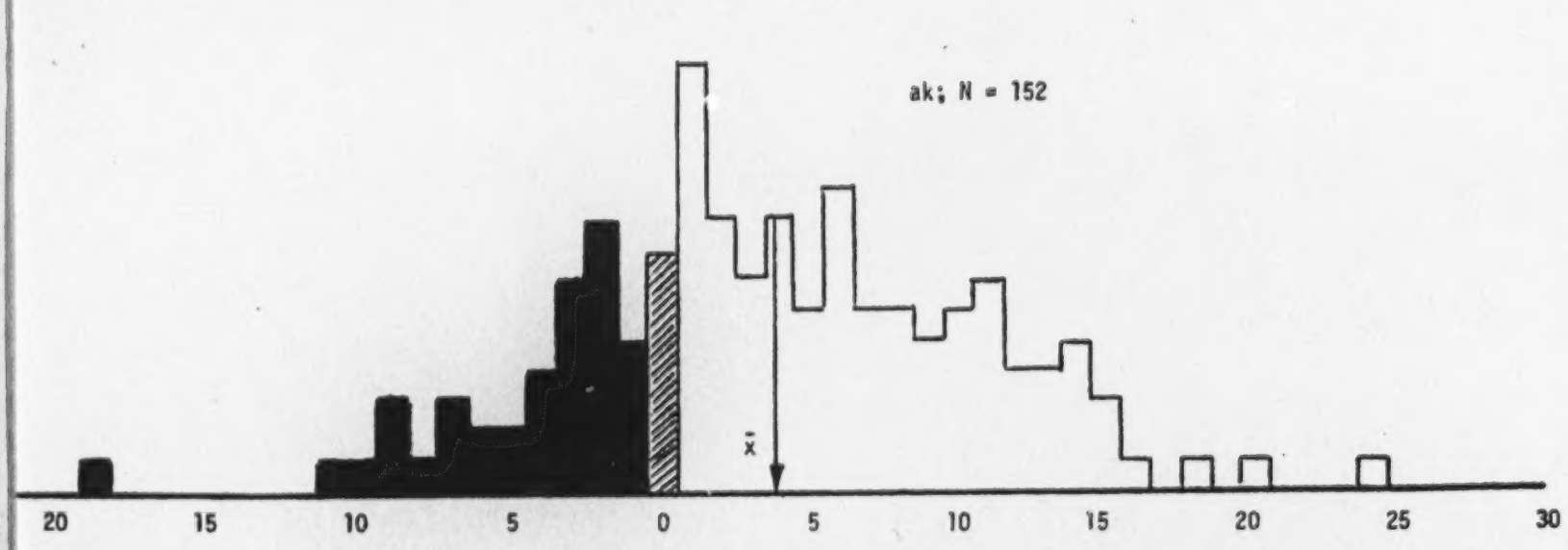
Table 18: Pre-test, Post-test, and Progress Means and Standard Deviations. sk; N = 235.

	Total			Boys			Girls		
	N	$\bar{x}$	s	N	$\bar{x}$	s	N	$\bar{x}$	s
Pre-test	235	59.11	14.83	127	58.24	15.01	108	60.13	14.61
Post-test	235	66.99	14.13	127	66.35	14.15	108	67.74	14.13
Progress	235	7.88	8.04	127	8.11	8.39	108	7.61	7.65

The progress made in sk is approximately the same as that made in grade 7 when the same grammatical structure, the passive voice, was taught (GUME 3). There is reason to believe that this progress is about the same as would be found after a year's teaching, though without the teacher's paying special attention to this specific structure (see p. 93 below). Thus, since progress is made by the pupils, there should be a fair chance for method differences, if any, to appear. A noteworthy fact considering the progress scores is that in all cases they are exceeded in size by the respective standard deviations. This means that the within-course variance is great and that a number of pupils make negative progress, i.e. they regress. This fact is also apparent from fig. 4 on the next page; the black field signifies regress scores. Incidentally, it is hardly probable that regress scores of 20 and 13 respectively (both are found in the figure) are true scores. Most likely they are test effects, caused by lack of motivation on the Post-test occasion. Similarly, very high

Fig. 4 : Distribution of Progress Scores for sk and ak.





progress scores may be explained as test effects because of low motivation on the pre-test occasion. However, all scores, whatever their nature may be in this respect, have been included in the analyses.

In the table below the corresponding values for ak are presented.

Table 19: Pre-test, Post-test, and Progress Means and Standard Deviations. ak; N = 152

	Total			Boys			Girls		
	N	$\bar{x}$	s	N	$\bar{x}$	s	N	$\bar{x}$	s
Pre-test	152	31.52	7.08	94	31.10	7.48	58	32.21	6.38
Post-test	152	35.24	8.45	94	35.98	8.96	58	34.05	7.47
Progress	152	3.72	6.80	94	4.88	7.21	58	1.84	5.64

The progress in ak is of the same magnitude as that found in GUME 3. However, the total progress is of very limited size and the within-course variance is great (see also fig. 4 on the preceding page). Nonetheless, it is theoretically possible for teaching method differences to exist. The boys exceed the girls in progress; the difference is significant at the 1 % level (cf corresponding differences in the case of scholastic aptitude on p. 72, table 15.

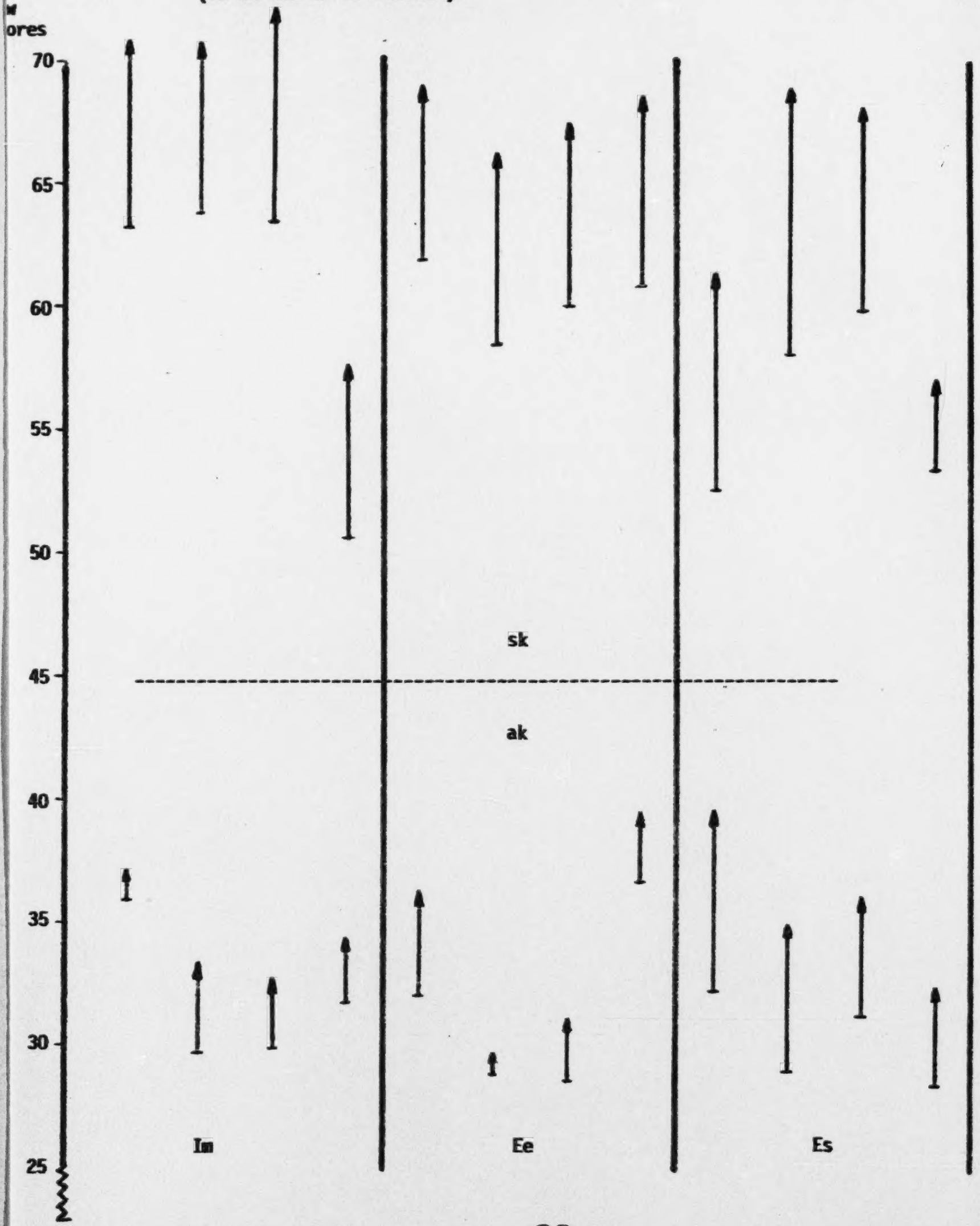
Progress - Main Effects.

In the present section a number of analyses will be made in order to find an answer to the question: which of the three methods proved to produce the best learning effects? However, before we proceed to the analyses, we shall discuss figure 5 (next page), intended to visualize the outcome of the study.

School class progress. In figure 5 the twenty-four school classes participating in the experiment are represented by arrows. The bottom end of each arrow signifies the pre-test score of one school class, the top end signifies the post-test score and the length of the arrow symbolizes the progress made. To the left the scale (raw scores) is indicated.

The perhaps most striking feature is the marked division into two groups of arrows, one at the top and one at the bottom of the figure. The two groups are sk and ak; at no point do sk and ak arrows overlap. As far as length of arrows is concerned there is also a great difference between sk and ak; the sk classes generally make significantly greater progress than the ak classes. These facts, pointed out in tables 18 and 19 , are very prominent in the figure. When the two groups of arrows are considered separately, the main impression becomes one of

Figure 5 : The Progress of the 24 Experimental Classes (12 sk and 12 ak Classes).



variation between school classes rather than one between methods (Im/Ee/Es); as a matter of fact, a number of classes have post-test scores which are lower than the pre-test scores of other classes. This strong variation between classes within courses is interesting per se. As regards length of arrows (= progress) the only pattern striking the eye is that the Es arrows in ak, the easier course, seem to be somewhat longer than the arrows in the Im and Ee methods.

Although the school class data give a levelled-out impression of the results, they should suffice to indicate the general tendency in the results. As it appears, there is no visible pattern in sk indicating priority of any method; in ak, progress seems to be somewhat larger in the Es group. However, since it is possible that the school class values may obscure individual data, we shall proceed to the latter.

Individual progress per method. The progress scores, i.e. the difference in raw scores between the Post-test and Pre-test scores, were analyzed with various covariates. One of these, the Standardized test in English is somewhat dubious as a covariate because it was given after the experiment and may theoretically be affected by the treatment. However, the teaching only concentrated on the passive voice, which is a very minor part of the course at large in ak and of limited scope in sk. We therefore accept the test as a covariate. The other covariates, which are all pure pre-experimental measures, are: DBA, PACT, and the Pre-test. In the following analyses the four covariates are treated separately and, in a final analysis, they are weighted together to a composite covariate. The sk and ak groups will be treated separately.

A word of caution is in order before presenting the various analyses. The reliability of a Progress score, calculated as a difference score between two measurements, is influenced by the reliabilities of the Pre- and Post-test as well as the correlation between the two. The reliability was calculated in both courses (as for the formula, see Ferguson, 1959, p. 285). In ak it proved to be .02 which means that most of the total variance of differences is error variance. The corresponding value for sk is .36. Thus the results as far as the progress measures are concerned should be treated with great caution in sk; in ak they are grossly unreliable. However, for the sake of completeness they have been included below.

Table 20: Analysis of Covariance, sk.
Dependent Variable: PROGRESS
Covariates: DBA, PACT, the Pre-test, the Standardized Test in English, and the Weighted Sum of the Four.

	Adjusted means				ss'_y			
Covariate:	Im	Ee	Es	F-ratio	Between	Within	df	b_w
DBA	7.53	7.53	9.04	.782	102	13854	2/210	.081
PACT	7.83	7.64	8.21	.095	13	14811	2/223	-.101
Pre-test	8.06	7.86	7.74	.034	4	13219	2/231	-.194
Std Test Eng	7.78	7.72	8.27	.100	13	14636	2/224	-.038
Total	8.07	7.48	8.62	.429	43	9926	2/199	

ss'_y = adjusted sum of squares in the dependent variable

b_w = the within-groups regression coefficient

It is a somewhat disconcerting fact that the correlations between the dependent variable and the different covariates are low or, in a majority of cases, negative; this is indicated by the within-groups regression coefficients (see the last column in the table). Thus no gain in precision is achieved by resorting to analyses of covariance; rather the contrary. However, for the sake of completeness, and also because the negative coefficients are interesting per se, we have presented the analyses. The rather high negative regression coefficient for the Pre-test is worthy of comment: the criterion test has obviously become too easy for the sk group; on the Pre-test they score 59.11 (see table 18), which is equal to having 62.9 % of the items correct before the experiment started. The distribution of scores on the Pre-test also shows that a fairly large number of sk pupils were near the maximum score. This ceiling effect is even more marked on the Post-test; thus the higher the pupils scored on the Pre-test, the smaller the probability for them to progress. The negative correlation between Progress and the Pre-test is thereby explained.

The negative regression coefficients initiated a digression. If we return to table 20 it becomes obvious that no teaching method has demonstrated any superiority over the others; all the F-ratios are

insignificant. This result is in line with the GUME findings so far. In the table below the corresponding analyses for ak are given.

Table 21: Analysis of Covariance, ak.
Dependent Variable: PROGRESS
Covariates: DBA, PACT, the Pre-test, the Standardized Test in English, and the Weighted Sum of the Four.

	Adjusted means				ss'_y			
Covariate:	Im	Ee	Es	F-ratio	Between	Within	df	b_w
DBA	2.17	2.13	5.89	4.382	342	4869	2/116	.226
PACT	2.30	2.61	5.96	3.956	366	6606	2/135	.149
Pre-test	2.96	2.22	5.84	4.544	377	6509	2/148	-.236
Std Test Eng	2.50	2.61	5.96	4.603	367	6104	2/144	.146
Total	2.90	2.47	4.24	.786	49	3443	2/108	

In the case of ak there is a clear tendency for one method to be ahead of the others, namely the Explicit-Swedish. All the F-ratios for the covariates treated separately are significant at the 5 % level; when the covariates are added together to a composite measure, however, the significance disappears. As in the case of sk, the dependent variable correlates negatively with the Pre-test. Although no ceiling effect is discernible for ak on the Pre-test and Post-test, it is obvious that for some reason the high scores on the Pre-test are generally those who made lesser progress.

In the following table two complementary analyses of covariance are given, one for sk and one for ak. In both cases the Post-test is the dependent variable and the Pre-test the covariate.

Table 22: Analyses of Covariance, sk and ak.
Dependent Variable: POST-TEST
Covariate: Pre-test

	Adjusted means				ss'_y			
Course	Im	Ee	Es	F-ratio	Between	Within	df	b_w
sk	67.17	66.96	66.84	.037	4	13219	2/231	.806
ak	34.48	33.74	37.36	4.544	377	6509	2/148	.765

The correlation between the dependent variable and the covariate is high in both sk and ak; thus the precision is increased considerably in this analysis as compared to an analysis of variance. In sk the results from the previous analyses are confirmed; there are no significant differences between the teaching methods. In ak the tendency remains for Es to be ahead of the two other methods. It is thus possible that the kind of explanations used in our Es method has had a facilitating effect on learning. However interesting this tendency is we prefer to interpret it with the greatest caution for the following reasons:

a) Progress is generally so low in ak as to make differences between methods of limited interest

b) When the different covariates were weighted together, the significance disappeared

c) In ak the reliability of the Progress score was extremely low.

Progress - Interaction.

The present investigation is concerned with two different populations; it has been shown clearly that they are quite distinct groups in all kinds of cognitive variables. All computations hitherto have been performed separately for the two courses. However, since all pupils were tested for scholastic aptitude by the same test (DBA), we shall tentatively perform an analysis of variance (two-way) with the total group (sk + ak) divided into three groups (Upper, Middle, Lower) according to DBA-scores. The procedure is admittedly somewhat inadequate since two distinct populations, having read two separate courses for three terms, are put together in the analyses. However, it may be excused as a means of finding out if any interaction exists between teaching method and ability level in the experimental population at large. The following table makes clear how the sk and ak groups were distributed on ability levels.

Table 23: Distribution of sk and ak Pupils on Three Levels of Scholastic Aptitude According to the DBA Test.

DBA level	sk	ak	Total
Upper	109	7	116
Middle	66	33	99
Lower	39	80	119
Total	214	120	334

In the following table the analysis of variance (two-way) is given.

Table 24: Analysis of Variance (Two-Way); sk + ak.
Dependent Variable: PROGRESS

Ability level	Teaching method Im	Ee	Es	Total:
U	7.07	9.80	9.97	8.86
	(42)	(41)	(33)	(116)
M	6.23	5.11	9.58	6.86
	(31)	(37)	(31)	(99)
L	2.83	2.36	5.08	3.34
	(35)	(47)	(37)	(119)
Total	5.45	5.61	8.06	6.30
	(108)	(125)	(101)	(334)

Source of variation	Sum of squares	df	Variance estimate
Rows (U, M, L)	1832	2	916
Columns (Im, Ee, Es)	500	2	250
Interaction	233	4	58
Within cells	18487	325	57
Total	21052	333	

$F_i = 1.026$ $F_c = 4.395$ $F_r = 16.101$

i = interaction
c = columns
r = rows

The interaction term is insignificant; thus in the total group no specific method proves superior at one level of ability simultaneously as another method is the best at another level. On the contrary, there is a significant column effect, i.e. one method, again the Explicit-Swedish, tends to be the best irrespective of ability level. These results appear to confirm those found in our previous analyses.

School Class Data Analysed.

It may be argued that the school class mean is the proper unit of analysis in a study like the present one (Wiley, 1969, p. 213). In the Pennsylvania project (see above, p. 6 ff) all method comparisons were based on that unit. Some of the analyses in GUME 4 (Lindblad & Levin, 1970, p. 77 ff) were also based on the school class means. In the present investigation the total population consists of two separate groups, each containing 12 classes. If, in this case, an analysis is performed on school class means, the number of observations becomes very limited. According to Carroll, for instance, the number of observations in each treatment should amount to 20 for conclusions to be valid (Carroll, 1969, p. 216). However subjective this judgment may be, it is obvious that in our case the limited number of cases only justifies very tentative conclusions. Since a tendency was found in ak for one method to excel the others, an analysis of covariance was computed with the ak school class mean as the unit of analysis. The following table gives the observations which the analysis was based on.

Table 25: School Class Means (ak) for the Pre-test and the Post-test. N = 12

	Im		Ee		Es	
	Pre-test	Post-test	Pre-test	Post-test	Pre-test	Post-test
	36.00	37.23	32.13	36.38	32.32	39.74
	29.67	33.42	28.93	29.71	29.07	35.13
	35.18	37.82	28.72	31.22	31.42	36.25
	31.79	34.57	36.67	39.69	28.43	32.71
$\bar{x}$:	33.16	35.76	31.61	34.25	30.31	35.96

The result of the analysis is presented in the next table.

Table 26: Analysis of Covariance of School Class Means (N = 12).
Dependent Variable: the POST-TEST
Covariate: the Pre-test

Sources	df	ss_x	sp	ss_y	$ss_{\bar{y}}$	df	$ms_{\bar{y}}$
A	2	16.27	- 0.59	6.99	23.34	2	11.62
w	9	77.95	82.90	102.63	14.47	8	1.81
Total	11	94.22	82.31	109.62	37.71	10	

$F = 6.42 \quad p < .05$

(Symbols as in Lindquist, 1953)

Even when the analysis (in ak) is undertaken at the school class level, the tendency for the Explicit-Swedish method to be the most efficient one prevails.

Additional Measures of Progress.

All analyses so far have been performed on raw scores. However, other measures may be more sensitive to progress (see p. 62 above). The following four analyses of variance, two per course, indicate whether the particular measures used may give any information not contained in the raw scores. Since the number of observations is limited, the results must be interpreted with caution. The first analysis is made on: Actual improvement x 100 / Possible improvement. Before each analysis of variance the actual school class means are presented.

Table 27: School Class Means for: Actual Improvement x 100 /
/ Possible Improvement; sk, N = 12

	Im	Ee	Es
	25.13	22.22	21.63
	23.10	22.31	26.38
	29.04	22.27	25.31
	16.72	23.72	9.11
$\bar{x}$:	23.50	22.63	20.61

In the following table the analysis of variance for the values above is given.

Table 28: Analysis of Variance (One-Way) of School Class Means on the Variable: Actual Improvement x 100 / /Possible Improvement; sk, N = 12

Source of variation	Sum of sqs	df	Variance estimate	F-ratio
Between	17.56	2	8.78	
Within	269.74	9	29.97	
Total	287.30	11		.29

The F-ratio is far from significant. Thus the tendency within the sk course from previous analyses persists.

In table 29 the 12 observations for the ak classes are presented.

Table 29: School Class Means for: Actual Improvement x 100 / / Possible Improvement, ak, N = 12

	Im	Ee	Es
	2.12	6.87	12.03
	5.83	1.21	9.35
	4.49	3.83	7.71
	4.48	5.23	6.54
$\bar{x}$:	4.23	4.29	8.91

Inspection of the table gives the impression that the Es group is still ahead of the other two. In the following table the result of the analysis of variance is given.

Table 30: Analysis of Variance (One-Way) of School Class Means on the Variable: Actual Improvement x 100 / Possible Improvement, ak, N = 12

Source of variation	Sum of sqs	df	Variance estimate	F-ratio
Between	57.68	2	28.84	
Within	41.37	9	4.60	
Total	99.05	11		6.27

The F-ratio is significant at the 5 % level; again, the priority of Es over the two other methods is confirmed.

The second measure used is: Progress x 100 / Pre-test. This measure gives comparatively great credit to progress scores for classes with low initial (= Pre-test) scores. The values for the sk classes are given below.

Table 31: School Class Means on the Variable: Progress x 100 / / Pre-test; sk, N = 12.

	Im	Ee	Es
	12.20	11.47	16.97
	11.17	13.54	20.79
	14.02	12.54	14.46
	14.39	12.91	6.90
$\bar{x}$:	12.95	12.62	14.78

Again, the corresponding analysis of variance is presented separately; see below.

Table 32: Analysis of Variance (One-Way) of School Class Means on the Variable: Progress x 100 / Pre-test; sk, N = 12.

Source of variation	Sum of sqs	df	Variance estimate	F-ratio
Between	10.84	2	5.42	
Within	112.33	9	12.48	
Total	123.17	11		.43

Nor did this measure give evidence of differences between the three teaching methods within the sk course.

In the following table the corresponding values for ak are given.

Table 33: School Class Means on the Variable: Progress x 100 / / Pre-test; ak, N = 12.

	Im	Ee	Es
	3.42	13.23	22.96
	12.64	2.73	20.88
	7.50	8.71	15.37
	8.78	8.18	15.09
$\bar{x}$:	8.09	8.21	18.58

When this measure is used, the priority of Es over the two other methods becomes even more pronounced. In the following table the analysis of variance of the school class means is given.

Table 34: Analysis of Variance (One-Way Classification) of School Class Means on the Variable: Progress x 100 / Pre-test; ak, N = 12.

Source of variation	Sum of sqs	df	Variance estimate	F-ratio
Between	290.12	2	145.06	
Within	145.77	9	16.20	
Total	435.89	11		8.95

The F-ratio is significant at the 1 % level.

In sum: the main investigation (treatment comparisons) has shown that in the case of the advanced course in English, sk, all results are completely in line with those found earlier in the GUME project, namely that no differences are found between the three teaching methods. The pupils in the easier course, ak, progress significantly less than those in sk. However, it is in the easier course that a relatively clear tendency for one method to excel appears. The Explicit-Swedish method is ahead of the two other methods in almost all comparisons that are made, both at the individual and at the school class level. Although the results should be treated with great caution for reasons given above, they do indicate that a teaching method containing explanations in the pupils' own language tends to facilitate learning, at least in comparison with a method with no explanations or one with the explanations in the target language. We shall return to the discussion of these findings in a later chapter.

Results on the Different Part Tests.

The overall progress scores for sk and ak have been given earlier (see p. 74 ff above). In this section the part test scores will be presented as they relate to the total sample as well as the three teaching strategies. The two courses will be presented and discussed separately; in the following table the values for sk are given.

Table 35: Results on the Part Tests per Method; sk.

	Sk; N = 235		Im; N = 70		Ee; N = 92		Es; N = 73		% x)	Max score
Pre-test	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s		
1	6.43	2.39	6.41	2.39	6.45	2.31	6.41	2.53	58.5	11
2	6.74	3.18	6.99	3.11	6.92	3.17	6.27	3.26	67.4	10
3	6.76	2.33	6.87	2.01	7.26	2.10	6.03	2.69	75.1	9
4	28.91	5.54	29.26	6.16	29.37	5.39	28.00	5.05	72.3	40
5	6.61	3.39	7.07	2.98	6.51	3.48	6.29	3.62	47.2	14
6	3.66	2.89	4.09	2.94	3.71	2.68	3.19	3.06	36.6	10
Total	59.11	14.83	60.69	14.80	60.22	14.13	56.19	15.48	62.9	94
Post-test										
1	7.12	1.95	7.67	1.78	6.75	2.00	7.07	1.95	64.7	11
2	7.26	3.08	7.19	3.40	7.61	2.77	6.89	3.12	72.6	10
3	8.00	1.78	8.23	1.30	8.11	1.91	7.66	1.97	88.9	9
4	31.21	5.98	31.43	6.44	31.58	6.04	30.53	5.46	78.0	40
5	8.56	3.21	8.66	2.67	8.78	3.19	8.18	3.69	61.1	14
6	4.83	2.90	5.27	2.71	5.03	2.83	4.16	3.09	48.3	10
Total	66.99	14.13	68.44	13.37	67.86	14.09	64.49	14.73	71.3	94
Progress										
1	.70	2.17	1.26	2.22	.30	2.12	.66	2.12	6.2	11
2	.52	3.08	.20	2.95	.68	2.63	.62	3.69	5.2	10
3	1.24	1.84	1.36	1.72	.85	2.04	1.63	1.59	13.8	9
4	2.30	4.81	2.17	4.49	2.21	5.22	2.53	4.62	5.7	40
5	1.95	2.50	1.59	2.36	2.27	2.47	1.89	2.63	13.9	14
6	1.17	2.20	1.19	2.19	1.33	2.40	.97	1.96	11.7	10
Total	7.88	8.04	7.76	7.14	7.64	8.80	8.30	7.96	8.4	94

x) The %-figures refer to the total mean for all pupils in relation to the possible number of items correct per test; the progress figures are the differences in per cent for the post- and pre-tests.

As might be expected, part test number 6, which is all written and requires the pupils to passivize active sentences, is the most difficult one (36.6 % correct on the Pre-test). Also on the Post-test occasion this part test ranks last; however, the progress score on it is not particularly low as compared to the others. Part test 2, which is also testing the pupils' ability to perform transformations, though from the passive into the active, is a great deal easier. This is partly due to the fact that the test is more passive than No. 6 (see Appendix A), but partly also to the fact that test No. 2 tends to be of the all-or-none type, i.e., if the pupil has understood the test principle, there is a good probability that he should manage all, or almost all, items. This is also indicated by the frequency distribution for part test 2. Part test 2 has the smallest progress of all the tests. The sk group performs particularly well on part test 3, the listening comprehension test. On the Post-test the sk pupils have 8 of the 9 possible items correct, which by far surpasses the result in the ak group (cf table 36).

In table 35 there is no evidence of differences between teaching methods on the part test level. An inspection of the table makes it clear that the relation between the methods, i.e. no differences, is the same on all part tests.

In table 36 on the following page the corresponding figures for the ak group are given.

As the table shows, the written test, no. 6,is definitely the most difficult in ak. The pupils can not, on their own, produce written transformations from the active into the passive. This inability is not only noticeable in the Pre-test but also equally striking in the Post-test. Part test 3 is a 4-choice and test 4 a 2-choice test. Considering this fact it becomes evident that, on the Pre-test, the values for these part tests are very near the chance level.

Concerning method differences on the part tests, the tendency for Es to excel prevails; thus there is no single part test that shows a pattern completely different from the total. The superiority of Es over the other methods seems to be particularly pronounced on part tests 2 and 4; as it happens, these two parts display comparatively little progress (3.2 % and .60 % respectively). In fact, the Im pupils have regressed on these tests. Since a regress (as an average) is hardly

Table 36: Results on the Part Tests per Method; ak.

	Ak; N = 152		Im; N = 50		Ee; N = 49		Es; N = 53			
Pre-test	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	% x)	Max score
1	1.51	1.64	1.72	1.61	1.31	1.77	1.43	1.54	13.7	11
2	3.34	2.62	3.76	2.62	3.27	2.92	3.02	2.32	33.4	10
3	2.23	1.88	2.38	1.71	1.92	1.78	2.38	2.11	24.8	9
4	22.96	4.14	23.52	4.14	22.84	3.74	22.55	4.49	57.4	40
5	1.34	1.24	1.52	1.28	1.37	1.27	1.15	1.17	9.6	14
6	.14	.45	.16	.37	.10	.31	.15	.60	1.4	10
Total	31.52	7.08	33.12	6.72	30.80	7.01	30.68	7.35	33.5	94
Post-test										
1	2.89	2.01	3.50	1.69	2.24	2.04	2.92	2.10	26.3	11
2	3.66	2.65	3.46	2.70	3.45	2.59	4.06	2.66	36.6	10
3	2.99	2.67	3.32	2.77	2.41	2.35	3.21	2.81	33.2	9
4	23.21	4.54	22.54	4.56	23.00	4.47	24.04	4.53	58.0	40
5	2.17	1.71	2.36	1.78	1.94	1.84	2.21	1.51	15.5	14
6	.32	.72	.52	.86	.14	.41	.28	.77	3.2	10
Total	35.24	8.45	35.70	7.61	33.18	9.10	36.72	8.35	37.5	94
Progress										
1	1.39	2.02	1.72	1.97	.94	2.10	1.49	1.96	12.6	11
2	.32	2.80	-.30	3.00	.18	2.73	1.04	2.56	3.2	10
3	.76	2.08	.94	2.15	.49	2.16	.83	1.96	8.4	9
4	.25	5.26	-.98	4.96	.16	5.11	1.49	5.49	.60	40
5	.83	1.46	.84	1.35	.57	1.55	1.06	1.47	5.9	14
6	.18	.70	.36	.85	.04	.41	.13	.73	1.8	10
Total	3.72	6.80	2.58	6.77	2.39	5.94	6.04	7.07	4.0	94

x) The %-figures refer to the total mean for all pupils in relation to the possible number of items correct per test; the progress figures are the differences in per cent for the post- and pre-tests.

---

a true score, it may perhaps best be explained as a test effect, i.e. low motivation on the Pre-test occasion. The progress scores on these part tests are very low also for the Ee method (.18 and .16 respectively). It is thus the case that a good deal of the priority of the Es method in the ak course is explained by the results on part tests 2 and 4, ironically enough two tests where the progress was very limited.

This finding stresses the importance of interpreting the method differences found with the greatest caution.

In sum: The part tests give the same information as the total test concerning the relation between the three methods. In the case of the all written part test (no. 6) the ak pupils' scores are negligible. The superiority of the Explicit-English method can be traced to two part tests where the progress is generally low; the hypothesis is forwarded that the superiority of Es may partly be explained as a test effect.

Drop-outs.

The drop-outs that will be referred to here are the pupils (sk: N = 20, ak: N = 16) who were absent for more than one lesson. In order to find out whether the drop-outs deviate in any systematic way from the experimental group, a number of comparisons between the two groups were made. The results of the comparisons are presented in tables 37 and 38, with sk coming first.

Table 37: Means and Standard Deviations for the Experimental Population and the Drop-outs (sk).

	Population (= pupils present 5-6 lessons)			Drop-outs (= pupils absent > 1 lesson)			
	N	$\bar{x}$	s	N	$\bar{x}$	s	t
DBA total	214	55.24	8.80	17	56.29	9.84	-.43
Grades total	233	29.95	7.42	20	27.45	6.25	1.69
Std test	228	49.41	12.81	16	49.25	12.80	.05
PACT	227	50.19	3.10	17	49.41	4.30	.74
Pre-test	235	59.11	14.83	20	56.75	16.05	.63
Post-test	235	66.99	14.13	20	59.10	18.82	1.83
Progress	235	7.88	8.04	20	2.35	9.17	2.61
Pupil Att.	200	22.55	4.43	13	21.77	3.72	.72
Absence	235	0.17	0.38	21	2.43	0.81	

We have included Absence as a variable in order to give an idea of the actual difference, in lesson time, between the two groups; the difference is of course highly significant. If we disregard this variable, the only significant difference appearing among the rest is that for Progress during the experiment, where those who attended

almost all lessons excel those who were absent two and a half lesson on the average. One conclusion is that the drop-outs did not deviate from the population in any background variable (it might have been hypothesized that the drop-outs have low scholastic aptitude, grades, etc.). Another conclusion is that the pupils profited from being present during the lessons in that they progressed relatively more. In the next table the corresponding values for ak are given.

Table 38: Means and Standard Deviations for the Experimental Population and the Drop-outs (ak).

	Population (= pupils present 5-6 lessons)			Drop-outs (= pupils absent 1 lesson)			
	N	$\bar{x}$	s	N	$\bar{x}$	s	t
DBA total	120	43.37	7.91	7	45.00	5.35	-.76
Grades total	148	23.43	5.50	16	22.69	4.38	.63
Std test	148	39.18	10.49	16	40.63	6.47	-.79
PACT	139	43.19	6.72	14	44.50	4.31	- 1.02
Pre-test	152	31.52	7.08	16	30.88	7.43	.33
Post-test	152	35.24	8.45	16	34.31	8.13	.43
Progress	152	3.72	6.80	16	3.44	8.18	.13
Pupil Att.	140	21.34	4.38	11	22.73	4.86	-.91
Absence	152	0.26	0.44	16	2.50	0.89	

Nor are there any significant differences in ak between the drop-outs and the experimental population as far as background variables are concerned; thus there appears to be no selection mechanism causing absence. However, the pupils who attend the lessons regularly do not progress more than the drop-outs.

Progress from a Different Point of View.

It has been stated earlier that the GUME 3 and GUME 5 projects have much in common. When GUME 3 was carried out in 1969 no control groups were used; nor was this the case in the present study. There were several reasons for not using control classes: The main purpose is to investigate differential progress between various treatment groups and not to investigate amount of progress. It is highly improbable that pupils who have not been exposed to concentrated teaching of a particular grammatical structure would progress because of maturation, etc. Rather it is to be supposed that their progress would be close to zero if measured after a time interval as great as the duration of the lesson series. Nor would it have been of particular interest to use a design with one control group being given the Pre-test but no treatment and another control group being given the treatment but no Pre-test; this type of design, intended to find out whether the Pre-test has sensitized the subjects to the evaluation instrument, is important when amount of progress, rather than differential progress, is the main concern.

However, viewed from quite another angle, amount of progress might be of interest to the language teaching profession. How long would it take in "ordinary" teaching, as compared to the case when concentrated attention is given to one particular grammatical structure, to achieve the same progress? As a kind of check on this, a number of school classes were given the Criterion test used in GUME 3 at the beginning and at the end of grade 7. The test was administered in August 1969 and May 1970 and 12 sk and 6 ak classes took part. Only parts 1-4 (see Olsson, 1969) of the test were given.

The result of this control group study will be presented in greater detail in a forthcoming synopsis (spring 1971); here we shall only briefly mention the main outcome. Incidentally, in the preceding GUME study (Lindblad & Levin, 1970, p. 118 ff) a similar study was made; the main finding was that the pupils learnt as much in the six project lessons as they do otherwise in one year.

In GUME 3 (see Olsson, 1969, p. 53) the progress made on parts 1-4 of the Criterion test was 10.13 points in sk and 4.16 in ak. In the follow-up study in grade 7 covering the school year of 1969-1970 the corresponding progress scores were 10.55 and 3.86 respectively. In other words, the earlier finding is confirmed, namely that the pupils learn as much in six lessons of concentrated teaching of one particular grammatical point as they do otherwise in one school year.

CORRELATION STUDIES

Product-moment correlations have been calculated between all variables in the two experimental samples, sk and ak. Compared to correlations accounted for in the GUME 4 experiment, the correlations in the two groups here are lower because of restriction of range in both groups. The correlation coefficients will be grouped in various ways and presented for sk and ak separately.

Intercorrelations between the Main Variables (the sk sample).

Table 39 on the next page gives the coefficients for sk. The following observations can be made:

Progress seems to be uncorrelated with most variables. An exception to this rule is the Pre-test, which correlates negatively with Progress. This is due to ceiling effects in the Pre-test; those who scored relatively high on the Pre-test thus progressed relatively little during the experiment. This fact has been commented on earlier (see page 79 above). Another exception is the Pupil Attitude variable, which shows a slight positive correlation with Progress. The same observation was made in GUME 4; it is an open question whether the positive attitude towards the lessons on the part of some pupils caused the relatively high progress, or whether the pupils' feeling of success was the cause of a positive attitude.

Pupil attitudes are similarly uncorrelated with most other variables. Interesting exceptions are the negative correlations with the part tests as well as the total of the Standardized Test in English. It is indeed difficult to give a good reason why the pupils scoring relatively high on the national test should tend to be the most negative towards the experiment.

Social class correlates around .20 with grades, the national test and our achievement tests, but somewhat lower with DBA.

PACT, the listening comprehension test, has the highest correlation (.519) with the EA variable, i.e. the listening comprehension part of the national test. This might be taken as an indication that PACT contributes with some unique variance.

Table 39: Intercorrelations between the Main Variables; sk, N = 235

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1 Social class	.146	.050	.089	.143	.206	.145	.259	.211	.207	.277	.243	.266	.087	.244	.263	.006	.022	.118	.326	.016
2 DBA Verbal		.248	.136	.600	.506	.616	.307	.546	.618	.517	.505	.626	.457	.551	.553	-.030	.024	.350	.398	.097
3 DBA Inductive			.437	.770	.231	.226	.344	.307	.352	.229	.236	.294	.181	.276	.340	.096	.137	.171	.290	-.021
4 DBA Spatial				.772	.114	.111	.367	.239	.199	.231	.181	.217	.108	.179	.250	.113	.042	.014	-.041	-.088
5 DBA Total					.378	.409	.470	.488	.519	.446	.413	.515	.332	.453	.518	.085	.084	.229	.270	-.018
6 Grades English						.696	.508	.834	.624	.647	.514	.691	.480	.721	.683	-.130	-.167	.724	.647	-.051
7 Grades Sw							.530	.838	.561	.512	.415	.576	.324	.571	.557	-.076	-.094	.598	.514	.022
8 Grades Maths								.802	.448	.450	.369	.483	.215	.481	.529	.041	-.049	.533	.576	-.115
9 Grades Total									.637	.613	.507	.676	.377	.675	.672	-.067	-.094	.703	.684	-.056
10 Std Achievem. EL										.665	.677	.900	.494	.684	.690	-.049	-.155	.459	.515	-.036
11 Std Achievem. EM											.629	.885	.405	.658	.656	-.057	-.242	.435	.556	-.017
12 Std Achievem. EA												.841	.519	.627	.600	-.102	-.236	.386	.373	.003
13 Std Achievem. Total													.518	.748	.745	-.066	-.239	.497	.586	-.026
14 PACT														.481	.477	-.044	-.057	.281	.124	-.022
15 Pre-test															.845	-.356	-.171	.597	.571	-.019
16 Post-test																.195	-.094	.532	.667	-.076
17 Progress																	.158	-.148	.051	-.098
18 Pupil Attitude																		-.077	-.277	-.130
19 Grades German x)																			--	-.014
20 Grades French xx)																				-.026
21 Absence																				

x) N = 143 xx) N = 64

The DBA variables show a very persistent pattern of correlations with grades, the variables of the standardized English test and our Pre- and Post-tests; the verbal part correlates around .55, the inductive part around .30, and the spatial part around .20 with them. This pattern has been found in the GUME studies so far (although the coefficients were higher where the pupil sample was not split in two courses).

The grades display a correlation pattern similar to the one just pointed out for DBA. Grades English correlate highest with the national test variables, PACT, and our Pre- and Post-test (around .60), Grades Swedish follow (around .50), and Grades Maths have the lowest correlation (around .40). Grades total correlate .60 on the average with the above mentioned measures; thus it seems as if the pupils' general capacity for studies, as it is expressed in the three main subjects, is about as good a predictor of proficiency in English as are the Grades English.

Grades German and French show a similar pattern of correlations with other variables. Both correlate around .25 with DBA, around .65 to .70 with Grades English, around .70 with Grades total, around .50 to .55 with the Standardized Test in English, and around .60 with our Pre- and Post-test.

The correlations between the various part tests of the Pre- and Post-tests and other variables will be discussed presently. First, however, the correlation matrix for ak corresponding to the one just discussed for sk, will be presented.

Intercorrelations between the Main Variables (the ak sample).

The correlations for ak are found in table 40.

Progress correlates negatively with the Pre-test also in the case of the ak sample. The reason for this is not quite clear since there was no ceiling effect in the Pre-test for the ak group; one hypothesis is that a number of ak pupils were somehow disturbed by the testing procedure (on the Pre-test occasion) and that there was, therefore, a higher probability for them to show progress on the Post-test occasion in comparison with those who took the Pre-test "naturally". There are correlations in the order of .25 between Progress and the DBA variables as well as between Progress and the variables of the Standardized English Test.

Table 40: Intercorrelations between the Main Variables, ak, N = 152

	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1 Social class	.122	.091	.075	.124	.086	.008	.018	.029	.013	.135	.008	.047	.036	.240	.165	.049	.092	.264	-.055
2 DBA Verbal		.284	.249	.607	.290	.330	.363	.419	.522	.382	.417	.523	.389	.366	.491	.221	.091	-.150	.006
3 DBA Inductive			.373	.777	.014	.164	.301	.189	.287	.247	.286	.325	.178	.127	.272	.203	.073	-.020	.062
4 DBA Spatial				.792	-.056	.147	.339	.189	.172	.158	.211	.216	.123	.066	.276	.273	.110	-.204	-.033
5 DBA Total					.072	.264	.449	.331	.399	.333	.388	.445	.284	.222	.446	.318	.123	-.177	.014
6 Grades English						.407	.391	.784	.461	.300	.378	.449	.455	.425	.371	.022	-.078	.366	.112
7 Grades Sw							.418	.713	.449	.272	.333	.415	.225	.324	.371	.127	-.005	.145	.087
8 Grades Maths								.812	.486	.305	.428	.483	.233	.420	.386	.045	.061	.247	-.032
9 Grades Total									.612	.385	.495	.589	.396	.510	.488	.080	-.012	.354	.054
10 Std Achievem. EL										.565	.710	.900	.643	.477	.628	.283	-.030	.154	-.003
11 Std Achievem. EM											.519	.758	.456	.440	.487	.145	-.056	.110	.034
12 Std Achievem. EA												.899	.688	.399	.554	.274	-.019	.110	.023
13 Std Achievem. Total													.717	.506	.652	.283	-.037	.147	.019
14 PACT														.458	.532	.193	-.046	-.065	.081
15 Pre-test															.629	-.259	-.064	.126	.064
16 Post-test																.587	.054	.174	.018
17 Progress																	.134	.061	-.044
18 Pupil Attitude																		.217	-.217
19 Grades German x)																			-.175
20 Absence																			

x) N = 44

Pupil attitudes are uncorrelated with all other variables (a coefficient of .16 is required for it to deviate significantly from zero). The pupils' responses on the attitude test will be discussed below (p. 103 ff).

Social class is uncorrelated with almost all variables; the few deviations from this pattern may be interpreted as chance occurrences. The main explanation of this finding is that the variance of the Social Class variable is very small; the majority of the ak pupils fall in only one of the three categories.

PACT; again, this test shows the highest correlation (.688) with the EA test, i.e. the listening comprehension variable, of the Standardized English test. Incidentally, the same finding was made in GUME 4 (Lindblad & Levin, 1970, p. 94). There should be little doubt that PACT is measuring a specific area of English proficiency.

The DBA variables again demonstrate the same pattern of correlations with grades and various achievement variables in English; the correlations between the latter and DBA verbal, DBA inductive, and DBA spatial are around .40, .20, and .15 respectively.

The grades present an interesting picture in the easier course in that Grades Maths correlate as high as Grades English with the various English test variables (around .40 on the average for both). However, the best predictor of success on the English tests (Nos. 10-16 in the matrix) are Grades total, which correlate around .50 with those variables. These findings are partly explained by the fact that the ak group, as compared to the sk group, scores relatively lower on the verbal factor and in verbal school subjects than on the spatial factor and in Maths. Thus, since the ak pupils are relatively low scorers in verbal respects, and their interest in English not particularly great, it is not surprising that Grades total and Grades Maths correlate higher or the same as Grades English with the various English tests.

Grades German; the number of pupils included are only 44, which means that a correlation coefficient of .29 is required for it to deviate from zero. The only interpretable correlation with any other single variable is that with Grades English, which is .37.

Pre-test and Post-test Correlations (the sk sample).

In this section we are mainly concerned with the validity aspect of the Pre- and Post-test. The various parts of them have been correlated with Grades, the different variables of the Standardized Test in English, PACT, and DBA. The correlation coefficients are given in table 41.

A general impression is that, in sk, the same pattern of correlations prevails among the Pre-test and Post-test parts and totals. One exception to this is part 4 of the Pre-test, where the correlations tend to be lower than for other part tests; this is not the case in the Post-test.

The totals correlate around .70 with Grades English and around .75 with the total score on the Standardized Test in English, which testifies to the validity of the Pre- and Post-test. Of the different parts, the last two ones (5 and 6), tend to show the highest correlations with Grades and the national test. This tendency for tests of a productive character to correlate relatively higher, as compared to tests of the fixed response type, with different criteria, is found in earlier GUME investigations (see, for instance, Lindblad, 1969, p. 62).

Part 3 is a listening comprehension test. It is interesting to note that this test is the one that shows the highest correlation with PACT. Again, there is evidence that PACT is measuring a relatively specific ability.

Pre-test and Post-test Correlations (the ak sample).

The correlations for the easier course are given in table 42.

The immediate impression is that the coefficients for ak are substantially lower than those for sk in the previous table. In the main these differences are explained by the lower reliabilities of the Pre-test parts (see p. 50 above).

The tendency found in sk for parts 5 and 6 to correlate higher than the other tests with the criteria, is not traced in ak. On the contrary, the all written test, number 6, correlates very low with Grades and the Standardized test. In the ak group this test simply did not work; tests of a productive character tend to be very difficult among the less talented pupils.

Table 41: Pre-test and Post-test Correlations; sk N = 235

		Grades				Std Test in English					DBA			
		Eng	Sw	Maths	Tot	EL	EM	EA	Tot	PACT	Verb.	Ind.	Spat.	Tot.
Pre-test	1	.570	.426	.300	.495	.458	.490	.477	.534	.380	.404	.144	.149	.317
	2	.501	.445	.461	.547	.461	.376	.448	.480	.289	.370	.152	.154	.300
	3	.495	.376	.346	.469	.614	.582	.564	.666	.483	.424	.354	.116	.398
	4	.450	.345	.212	.381	.469	.384	.406	.480	.354	.360	.171	.093	.282
	5	.623	.483	.449	.586	.542	.611	.486	.627	.380	.485	.248	.159	.405
	6	.684	.560	.502	.660	.581	.618	.514	.656	.318	.493	.230	.172	.403
	total	.721	.571	.481	.675	.684	.658	.627	.748	.481	.551	.276	.179	.453
Post-test	1	.461	.329	.293	.411	.451	.530	.446	.548	.342	.430	.228	.143	.359
	2	.437	.369	.457	.486	.421	.382	.338	.428	.291	.240	.246	.218	.327
	3	.361	.360	.250	.368	.473	.426	.454	.513	.490	.439	.340	.137	.407
	4	.490	.407	.382	.485	.538	.455	.444	.558	.321	.449	.268	.197	.414
	5	.623	.508	.493	.611	.582	.612	.529	.653	.426	.467	.250	.190	.409
	6	.635	.478	.411	.580	.561	.554	.475	.614	.359	.458	.217	.197	.394
	total	.683	.557	.529	.672	.690	.656	.600	.745	.477	.553	.340	.250	.518

Table 42: Pre-test and Post-test Correlations; ak N = 152

		Grades				Std Test in English					DBA			
		Eng	Sw	Maths	Tot	EL	EM	EA	Tot	PACT	Verb.	Ind.	Spat.	Tot
Pre-test	1	.133	.222	.181	.226	.330	.262	.300	.349	.368	.245	.161	.084	.210
	2	.190	.264	.330	.323	.234	.148	.119	.194	.171	.274	.244	.008	.217
	3	.398	.247	.218	.378	.432	.388	.438	.490	.378	.187	.147	.102	.191
	4	.264	.088	.233	.272	.235	.279	.226	.281	.281	.166-	.137-	.030-	.026
	5	.336	.281	.283	.384	.316	.267	.191	.294	.259	.310	.207	.138	.275
	6	.085	.144	.204	.189	.154	.224	.086	.168	.035	.110	.068	.156	.153
	total	.425	.324	.420	.510	.477	.440	.399	.506	.458	.366	.127	.066	.222
Post-test	1	.300	.295	.300	.383	.469	.461	.444	.529	.434	.295	.081	.176	.234
	2	.140	.348	.312	.339	.232	.158	.117	.194	.115	.154	.256	.115	.239
	3	.356	.171	.227	.342	.427	.464	.462	.521	.489	.319	.137	.239	.303
	4	.129	.098	.134	.163	.394	.173	.349	.371	.266	.340	.091	.101	.216
	5	.321	.413	.325	.436	.392	.346	.298	.397	.401	.286	.245	.246	.341
	6	.146	.058	.128	.132	.191	.186	.194	.221	.246	.253	.288	.158	.310
	total	.371	.371	.386	.488	.628	.487	.554	.652	.532	.491	.272	.276	.446

Part 3, the listening comprehension test, again correlates substantially with PACT.

On the whole, however, the part test correlations are of moderate size and fluctuate somewhat; we abstain from further comments.

Pre-, Post-test- and Progress Correlations.

The Pre- and Post-test parts have been correlated with each other and with the Progress variables. The matrices for sk and ak will be found in Appendix F. Here only a few comments will be made.

In sk the three last parts, 4-6, correlate highest with the total; this holds for both the Pre-test and the Post-test. Thus the parts accounting for most of the variance are tests which are relatively productive in character; the high correlations between part 4 and the total are mainly due to the fact that it contains relatively many items (40) and that no correction for guessing was made. In ak this particular test, part 4, is also the test which correlates highest with the total. If, however, we disregard this part we find no other test that correlates higher than the rest with the total.

If we correlate part 1 of the Pre-test with part 1 of the Post-test, we find that this correlation is higher than the correlations between part 1 of the Pre-test and any other part of the Post-test; similarly for the correlation between parts 2 and 2, 3 and 3, etc. This pattern holds for all pairs of tests in the sk group and for all but one in the ak group, namely part 4, which is in line with the low K-R (21) reliability of the test (.43).

In both the samples, sk and ak, Progress 4 correlates highest with Progress total. Thus the pupils' total progress (or rather, their change from Pre- to Post-test, i.e., in some cases a regress) is primarily explained by a corresponding progress (change) in part 4. If we consider the fact that the ak pupils progressed a quarter of a point on this test (cf table 36), it becomes evident that it has been a failure in the easy course.

THE PUPIL ATTITUDE TEST

The questionnaire consisted of questions of the open answer type as well as questions with fixed response alternatives (see Appendix B). First we shall comment on the responses to the latter type of items. They are eight in all and cover various aspects of the experiment; added together they can be considered to reflect the pupils' general attitude towards the lesson series. In the table below the key to the different questions is given.

Table 43: Pupils' responses on the Questionnaire (Questions with fixed response alternatives). sk and ak.

Question No.	Key	sk N	sk $\bar{x}$	sk s	ak N	ak $\bar{x}$	ak s
4	Learnt much (5) - little (1)	201	2.79	1.06	141	2.99	1.06
5	Fun (5) - boring (1)	201	2.88	1.27	141	2.45	1.09
8	Time went fast (5) - slowly (1)	201	3.07	1.28	141	2.59	1.17
9	Felt much less tired (5) - much more tired (1)	201	3.03	.98	141	2.84	.92
11	Sound quality: very good (4) - very bad (1)	201	3.17	.66	141	3.07	.78
13	Oral exercises: very good (4) - very bad (1)	201	2.25	.77	141	2.31	.83
14	Written exercises: very good (4) - very bad (1)	201	2.70	.66	141	2.54	.81
15	Reading texts: very good (4) - very bad (1)	201	2.65	.77	141	2.60	.85
Total		201	22.55	4.43	141	21.34	4.38

The sk pupils have a more benevolent attitude towards the experiment than the ak pupils; this finding is not surprising. The difference between the two groups is statistically significant ($t = 2.50$). Most of the difference is explained by the responses to questions 5, 8, and 9; the ak pupils regard the lessons as boring and tedious. The sk group is fairly neutral to the experiment; on the average it is judged to be about the same as ordinary teaching - whatever that means. The sk pupils seem to have appreciated the written exercises, and both groups think that the sound quality of the tapes was good.

The school classes vary a great deal in attitude to the experiment. The means for the school classes given in Appendix F testify to this. In ak no tendency between the three teaching methods is discernible as far as attitudes are concerned. In sk the tendency for the Im classes to have a more positive attitude was investigated by analysis of variance. Before the analysis is presented, the class means for attitudes are given below (cf Appendix F):

Table 44: School Class Means for Attitudes to the Experiment; sk N = 12

	Im	Ee	Es
	23.11	20.19	22.39
	22.33	24.62	22.75
	26.47	20.17	18.55
	24.07	22.28	21.93
$\bar{x}$	24.00	21.82	21.41

The table below gives the analysis of variance performed on these data.

Table 45: Analysis of Variance of School Class Means for Attitudes to the Experiment.

Source of variation	Sum of sqs	df	Variance estimate	F-ratio
Between	15.52	2	7.76	
Within	34.32	9	3.81	
Total	49.83	11		2.04

The F-ratio is not significant. Thus the tendency towards a difference between teaching methods in the case of attitudes towards the lesson series may be explained as a chance occurrence.

So far the pupils´ responses on the questions with fixed response alternatives have been considered. In the following section comments will be made on the free answers.

To get a view of the pupils´ spontaneous reactions a sort of simple content analysis was made of the answers. Twelve of the school classes were used in this analysis; i. e. only classes with odd identification

numbers were chosen. Categories of answers were established pragmatically, i.e. as an answer appeared that did not fit into a previous category, a new one was formulated. Finally twenty-two categories of a positive kind and sixteen of a negative kind were established; however, some of them contained only a few cases. Below a survey of the most frequent categories will be given.

Item 2: What was good about the experiment? (it should be remembered that only 50 % of the experimental population was included; thus the figures may be doubled).

The songs and music was the most popular feature of the project (51)

No home work comes next (43); this answer is more frequently used in the sk group than in the ak group.

Learnt more than during ordinary teaching (23)

More fun, more change (19)

Possible to control oneself (13)

The funny stories (12).

Item 3: What was not so good about the experiment?

Dull, too slow, too long pauses (98). This type of answer is by far the most common; the answers are equally distributed among the sk and ak pupils.

Too much repetitions, harping (40)

Learnt nothing (12)

Just listening to a tape-recoreder; no teacher (12).

As in the case of the positive answers all the categories with few answers are excluded. If the number of spontaneous answers is considered it becomes clear that the negative responses outnumbered the positive ones. A few examples of both the varieties (responses to item No. 16: Further comments):

"The whole thing was too ridiculous for grade 8".

"I don't like the kind of school where you have to sit just talking to a machine".

"We could have learnt this in two hours. If everything should be going at this speed the comprehensive school will take twenty years to pass through".

"I think the GUME project was fun and useful".

"Welcome back".

It is of course difficult to evaluate the free answers reliably. It seems as if they constitute a colourful image of the responses to the multiple-choice questions. There is undoubtedly a negative bias in both types of responses, and this bias is more pronounced in the case of the easier course.

Attitudes towards the explanations. The pupils of the two explicit strategies were requested to answer item No. 12 concerning the attitude towards the explanations (see Appendix B). The distribution of answers is given in the table below.

Table 46: Distribution of Responses to Item No. 12 of the Questionnaire Concerning the Explanations.

	5	4	3	2	1
sk	19	71	64	5	3
ak	11	52	26	3	2

The means for the two groups are 3.60 and 3.71 respectively. Both the pupil groups thus think that the explanations somewhat facilitated understanding.

General interest in English. The pupils were asked to indicate their liking for the various school subjects (see the first page of the Pupil Attitude Test, Appendix B). Since the results are of interest only with respect to English we shall not give any detailed account of the findings. Suffice it to say that in sk as well as in ak English ranked first among the academic subjects. Before English, in rank order, Gymnastics, Handicraft, and Drawing are found.

THE TEACHER ATTITUDE TEST

Two of the teachers preferred not to fill in the questionnaire; both were teachers of sk classes. As it happened, one of the teachers taught one sk and one ak class, and in that case only one questionnaire was obtained. The following account thus reflects the opinions of 21 of the 23 teachers.

Since the teachers did not take a very active part in the teaching, which was of the pre-produced variety, there is no intention to relate the background information about the teachers to the progress (or attitudes) of the respective school classes. Accordingly we will not give any detailed account of the background of the teacher sample; suffice it to say that the group consisted of 14 women and 9 men aged as follows: 21-30: 3, 31-40: 11, 41-50: 6, 51-60: 0, 61- : 1. Ten of the teachers were elementary school teachers, five of whom had further academic training in the subject of English, of the remaining eleven teachers all had a degree in English.

The first part of the questionnaire asked for the teachers' views on foreign language teaching in general.

Item No. 7: Which method could be predicted as being the best one for various levels of pupil ability? The distribution of opinions are given below:

	Im	Ee	Es	
Upper	2	10	12	24
Middle	2	1	17	20
Lower	7	-	13	20
	11	11	42	64

Two of the teachers marked all methods at the Upper level, arguing that it is of little consequence which method is given to talented pupils; they will succeed anyway. In general, the teachers think that the Explicit-Swedish method is the most promising one. At the upper level of ability they seem to prefer an Explicit method; if the explanations are given in English or Swedish is of minor importance. According to some of the teachers, the Im method is the best one at the lower level.

Item No. 8: Which method do the teachers use themselves?

Im: 2 Ee: 5 Es: 14

The frequencies above are very natural against the background of the answers that the teachers gave in the previous item.

Item No. 9: Do the teachers think that the pupils at the upper stage of the comprehensive school (grades 7 through 9) should have a course book in grammar?

Yes: 16 No: 5

The general orientation of the teachers seems to be that a grammar book would further conscious control of grammatical structures.

Item No. 10: How often should explanations be given?

Each lesson: 2
Fairly regularly: 14
Every now and then: 5
Never: 0

The inclination of the teacher group is definitely towards giving explanations, although the amount and intensity may vary.

Item No. 11: When explanations are used, in what language should they be given?

Swedish: 17 English: 4

The tendency is in accordance with the responses to items 7 and 8.

Item No. 12: How should the explanations be given?
By the teacher: 11
By a pupil, though rounded off afterwards by the teacher: 10

The teaching procedure indicated by the second alternative is the one which corresponds to the Recommendations of Lgr 62 and Lgr 69 II:Eng. Nevertheless, the teacher opinions are equally distributed on the two alternatives.

Item No. 13: To what extent should English be spoken during the lessons in ak, the easier course?

The variation in opinions among the teachers is very great: the lowest value is 25 %, the highest 100 %. The mean for the 19 teachers who answered the question is: 69.7 %.

Item No. 14: To what extent should English be spoken during the lessons in sk, the advanced course?

The teachers are relatively unanimous as regards sk. 12 of the 19 teachers answer between 80 and 90 %, the values ranging from 70 % to 100 %, the mean being 85.1 %. One of the teachers who abstained from answering this item commented that no percentage can be given since the actual relation English talk/Swedish talk depends on a number of things, such as the teachers' own capacity, the pupils' interest and ambitions, etc.

In short: the bias of the teacher sample in our experiment seems to be towards the Explicit-Swedish method generally, although some express the possibility for the Implicit method to work at lower levels of ability. The teachers think that explanations should be given regularly and that the pupils will benefit from the use of a special course book in grammar. English should be spoken most of the time in both the courses, although a slight difference is noticeable between sk and ak (85 % and 70 % English respectively).

Part 2 of the questionnaire focused on the experiment proper.

In the following account of teacher responses to the second part of the questionnaire only a sample of the items will be treated. Items referring to general aspects will be included, whereas those concerning technical details, etc., are excluded. The answers given by the teachers of the advanced course are commented on first.

No. 1: What was methodologically good about the teaching method you happened to get?

The answers do not seem to be correlated with teaching method; however the teacher comments will be referred to method:

Im: "The pupils were activated, Good examples, The songs were very stimulating (the perhaps most frequent comment), Good switches from one exercise to another".

Ee: "Ample amount of examples, Logical progression from the simpler to the more complex forms of the passive structure, The lesson series was well-planned, The contents were varied".

Es: "Good changes between music and work, The general sequence was good".

No. 2: What was methodologically bad about the teaching method you happened to get?

Im: "The singular and plural forms of the verbs were only occasionally commented on; it ought to have been done more often and more markedly, The lessons contained many difficulties not related to the passive voice".

Ee: "The pupils reacted against taped lessons, The pupils thought chorus reading was ridiculous".

Es: "Much too long time devoted to a single grammatical problem, Too many lessons, Too long lessons".

No. 3 a: Comments on the explanations (the E groups only):

Ee: "Too fast, Control questions about the pupils' comprehension were almost completely lacking, The explanation contained some difficult vocabulary which caused confusion".

Es: "Dull, Unnecessarily lengthy, Acceptable, More explanations might have been supported by pictures, etc."

No. 3 b: Comments on the oral drills:

Im: "Good, but in my class they might have been a little more difficult, The pupils were a little irritated at the slow speed, They were well chosen".

Ee: "Often too little time for the pupils to find the correct answer, Too short pauses for the class to react, The time allotted to pupils' answers was too short; it would have been wise to repeat some of the questions".

Es: "Difficult to get chorus answers from the pupils, Mostly good, Sometimes too long sentences for the pupils to say after".

No. 3 c: Comments on the written exercises:

Im: "Good, Useful, although they might have been somewhat more difficult in my class, Very well chosen, Instructions clear, Good examples".

Ee: "Too long time for the written exercises, especially in sk, Good, but most pupils had finished writing fairly soon and then they were sitting there just inactive".

Es: "Too long pauses after the written exercises, Some pupils had difficulty in correcting mistakes because of limited time".

No. 3 d: Comments on the reading exercises:

Im: "Interesting and well chosen, Filled with good examples of the problem at issue, Nice texts, Good".

Ee: "Some enjoyed the texts, some thought them sissy, Good reading texts".

Es: "The text aroused the pupils´ interest, Good, When the tape says 'Read after me' the written phrases should have been underlined in order for the pupils to find them more easily".

It is not easy to find a clearcut pattern through the answers given thus far. Our general impression, however, is that the positive and negative comments balance each other relatively well, that the written and reading exercises seem to be the parts most readily accepted, that the oral exercises were delivered at a too high speed and that the explanations tended to be slightly boring.

No. 4: The tempo during the lessons. Here we shall not quote the responses. The general tendency from the earlier questions is found again; the oral exercises did not always give the pupils enough time for responding, whereas the written exercises and the explanations tended to be lengthy.

No. 5: The sound quality of the tapes was generally considered good or very good. Single negative comments may be due to the listening conditions prevailing in a particular classroom.

No. 6 concerns the teachers´ opinions about pupil interest, discipline and learning effects in comparison with ordinary teaching. Most comments indicate that the students´ interest flagged as the lesson series passed by, and that the discipline was not affected in any direction by the experiment. Practically all the teachers abstained from speculating on the learning effects.

No. 7: Comments on the Pre- and Post-test were generally accepting or positive.

No. 10: Most teachers were of the opinion that "their" method would work. The following suggestions for improvements or alterations were obtained:

The Implicit method with explanations added would be the method, according to one teacher. Shorter work sessions are recommended by some, and a shorter lesson series is considered to be sufficient for this particular grammatical point.

No. 11: Further comments:

Im: "It would be of great value if the teacher had pre-produced materials of this kind at his disposal; however, the teacher should be allowed to make his own comments and explanations, The Im method does not give the pupils a clear enough view of actives and passives, The teacher should be allowed to operate the tape-recorder on his own since he knows when all the pupils have finished an exercise, Sometimes the lessons intruded upon the following ten-minutes break".

Ee: "I was a bureaucrat with bundles of papers rather than a teacher; it was frustrating to have so little contact with the pupils".

Es: "The pupils need opportunities for questions and comments, My class have had four different teachers this year, so the experiment was looked upon sceptically" (the "fifth teacher" ?).

Below the corresponding comments for the teachers of the easier course will be given.

No. 1: What was methodologically good about the teaching method you were assigned?

Im: "Good listening training and training in chorus reading, It was good for the pupil to hear the right answer immediately after his own response, The oral exercises where the pupils responded according to a certain pattern were good, The lessons contained a variety of exercises".

Ee: "The pupils are looking forward to a 'real' lesson again (!), The pupils are more or less compelled to understand instructions in English, Many examples, The oral exercises".

Es: "That they got the explanations in Swedish - otherwise it would not have worked, The change in activities".

No. 2: What was methodologically bad about the teaching method you happened to get?

Im: "That it contained no explanations, It was difficult to squeeze the lessons into an ordinary teaching period - in ak the trivialities of handing out papers, urging the pupils to take off unnecessary clothes etc., are comparatively time-consuming; with a fixed teaching program the consequence is that the forthcoming break is intruded upon, which causes irritation, Too short pauses".

Ee: "Many thought it was very difficult and became inattentive after a while, The pupils in ak can not stand thinking and listening for such long periods at a time".

Es: "The switches from one tense to another were sometimes confusing, and the pupils did not understand which tense was the expected one, It should have been made clear whether what was said on the tape was also to be found in the lesson materials for the pupils".

No. 3 a: Comments on <u>the explanations</u> (only the Explicit groups):

Ee: "Difficult; they do not know what grammatical words (!) are, Some of the explanations are difficult to understand, at least when no complementary explanation is given in Swedish, Should be given in Swedish in the ak course, Much too complicated, Caused bewilderness".

Es: "They were good, on the whole, Very good, although the less talented pupils often found them lengthy and could not concentrate, Too little time for the red pages (pages where passive and active sentences are contrasted)".

No. 3 b: Comments on <u>the oral exercises</u>:

Im: "Valuable, I think, but the pupils did not like the purely oral drills; they would have liked a picture or a piece of text as support, It would have been fine if one or two examples had been given in written form, Very good".

Ee: "Fairly good, Did not work well since the pupils did not always understand what was said on the tape, When the exercises contained words unfamiliar to the pupils, there was complete silence, Good".

Es: "The pupils were often insecure about what to answer, especially at the beginning of an exercise; however, the exercises proper were excellent, The so-called free exercises are not free because the time is insufficient; there was not enough time to elicit individual answers".

No. 3 c: Comments on the written exercises:

Im: "Too monotonous, the pupils were bored, Much too difficult for ak pupils, They were good; unfortunately the group had a tendency to look at the correct answers in advance, Rather difficult for the majority of the pupils in ak".

Ee: "Too long time devoted to written exercises, The exercises were good as such, but the pupils would have preferred them to be less numerous, Too difficult for the majority, The correct answers after the exercises were delivered too quickly".

Es: "Too difficult for ak, They were very good, too; the exercises were excellent and no problems occurred, When the right answers were given they ought to have been commented on; many faults remained uncorrected".

No. 3 d: Comments on the reading exercises:

Im: "Good exercises in chorus reading, Some were excellent, for instance 'Murder at Nightshade Hall'; unfortunately the group was not equally interested in all the reading exercises, Rather difficult in ak".

Ee: "Too complex, Most pupils understood them and were amused by them, Good, they aroused interest".

Es: "The pupils were amused by them; I don't think the pupils realized that they contained examples of the passive voice, however, Probably the exercises would have given more if there had been questions on the content".

Although it is hazardous to summarize the teacher comments thus far, the general impression is slightly towards the negative. Many of the exercises are said to be good, even excellent, "as such", but they seem to have been above the ak level mostly. The most positive comments concern the listening aspect of the lessons.

No. 4: The tempo of the lessons. Here we shall not quote the teachers but only summarize their opinion: The pauses in connection with oral drills were too short and the duration of the written exercises was too long. The tempo during the oral exercises seems to have been satisfactory.

No. 5: The sound quality of the tapes was considered good by the teachers.

No. 6: What were the teachers' opinions about pupil interest, discipline, and learning effects in comparison with ordinary teaching? Although there are teachers who hold that the motivation on the part of the pupils was greater during the project lessons than during more ordinary kinds of teaching, the majority of opinions are biassed towards the negative. Discipline seemed to be about the same as during the conventional lessons. In general, the teachers desist from predicting learning outcomes.

No. 7: Comment on the Pre- and Post-test: Some of the teachers suspect that the Pre-test, because of its great difficulty, caused a negative attitude towards the experiment.

No. 10: The teachers were asked to suggest improvements on "their" method. The spontaneous answers given indicate that the Im method with explanations added to it would produce an optimal method. Shorter work periods and a shorter lesson series for a grammatical point of the frequency of the passives are suggested.

No. 11: Further comments:

Im: "In ak it is imperative that the teacher should have close contact with his pupils; teaching of this kind does not take this fact into consideration, If I had known in advance how much extra work this project entailed I would not have agreed to participate, Was it necessary to include all tenses?".

We have tried not be biassed in our sampling of the teachers' answers and comments. The general impression of the comments from the ak teachers is definitely one of negative bias. As was mentioned early in this report (p. 45 above) the teaching materials is a compromise between what may be optimal for sk and ak respectively. The teacher comments confirm our own view that the lesson contents is more in line with sk than with ak standards.

The second part of the questionnaire contained one question with fixed response alternatives:

On the whole I think that the time set aside for the experiment has been:

(1) almost completely wasted
(2) rather ineffective
(3) approximately as usual (with respect to learning)
(4) fairly effectively used
(5) very effectively used

The teacher responses are given below.

	Response alternative 5	4	3	2	1	
sk	0	4	2	2	1	9
ak	2	5	3	2		12

One sk teacher abstained from answering. The frequencies in the table above give the following means: sk: 3.00 ak: 3.58. Surprisingly enough, the rather negative attitude on the part of the ak teachers, as expressed in the open answers, is not reflected in this particular item.

DISCUSSION OF RESULTS

The main results obtained in the advanced course, sk, are in complete accordance with earlier findings in the GUME project; no differences in learning effects were evidenced. The three teaching strategies, Im/Ee/Es, brought about as much, or as little, learning of the passive voice. However, in the easier course, ak, there was a tendency for the Explicit-Swedish method to surpass the two others. The Es method proved superior when the analyses were performed at the individual as well as the school class level and when various measures of progress were used. However, for the results to be treated with the caution they deserve, the following facts should be remembered:

a) Progress was generally very low in the easier course, and this may in itself put a limit to the interest of the foreign language teaching profession in method differences. To a great extent the small progress in ak was due to the fact that the teaching materials presented was a compromise between what might be considered optimal in each course and apparently the difficulty level gravitated more towards the sk than the ak standards. Evidence of this was also obtained from the attitude test. Although the teaching materials may be accepted for testing the main hypothesis of the experiment - whether explanations facilitate learning - it is clear that the teaching program did not function well in ak. To what extent this is due to the inherent difficulty of the materials or the fact that no modifications, i.e. concerning speed, were made in the program is difficult to say.

b) The Progress scores in the easier course were grossly unreliable.

c) The superiority of the Explicit-Swedish method was most prominent in part tests 2 and 4, i.e. the two part tests where the easier course generally made very little progress. In fact the Im pupils regressed on both these part tests; since a regress of this kind is hardly a true score, it may be interpreted as a test effect, probably low motivation on the Post-test occasion. The progress of the Ee pupils was almost equally insignificant on these two part tests; .18 and .16 respectively.

Obviously there are a number of reasons why the results favouring one particular method in the easier course, should be interpreted with care.

However, considering that the tendency for Es to surpass the other methods was consistent throughout the analyses, whether they were performed at the individual or the school class level and no matter what measure of progress was used, the results should not be dismissed as completely irrelevant. Rather we would prefer to advance the following hypothesis - which needs further testing: Explanations of the kind used in the present study, i.e. fairly simple observations, in the pupils' mother tongue, on how the different parts of the sentence behave, facilitate learning at the lower levels of ability.

It is tempting to compare the results of the present study with those of GUME 3, partly because the same grammatical structure was taught in both the experiments and partly because the same methodological expert was the person mainly responsible for the construction of teaching materials in the two studies. In GUME 3 (see Levin, 1969, pp. 64-65) there was also a tendency for the Es method to excel; however, at that time three parallel studies were performed and the tendency just mentioned happened to conflict with an opposite tendency (Im>) in another part study. The tendency in GUME 3 (Es>) was furthermore found at all ability levels (upper, middle, lower). It is hardly probable that the particular grammatical structure, the passive voice, should lend itself more readily to explanations (as opposed to non-explanations) than other structures. Nor is it probable - although the possibility cannot be completely ruled out - that bias on the part of the constructor of the teaching materials has contributed to the method differences found. A third and more probable explanation would be that the kind of explanations used in GUME 3 and GUME 5 differed somewhat from those used in the other part studies. It is the contention of the present authors that the explanations in GUME 3 and GUME 5 were comparatively "simpler" than those of the other part studies, where the grammatical points were also attacked from a semantic and/or a functional point of view. However, it would take a meticulous content analysis to prove this. A negative interpretation of the findings in GUME 3 and GUME 5 would of course be that in those studies the Im and Ee methods were comparatively bad as teaching

procedures. However, here is not the place to make comparisons between the results of the various GUME projects; this will be done in a forthcoming synopsis (spring 1971).

In the perspective of the current debate in Sweden and elsewhere on foreign language teaching methodology, the present investigation seems to have given as enlightening results as the earlier GUME studies. No dramatic differences, clearly favouring one method, have been found. The tendency for one method to be superior in the easier course is hazardous to interpret but food for thought and hypotheses to be tested.

SUMMARY

The present investigation is a direct continuation of earlier GUME studies. Since these produced non-significant differences in an assessment of three teaching methods compared, it was considered worthwhile to perform a new experiment with modifications that might increase the probability of detecting true differences, if such existed, between methods.

This part project, GUME 5, and GUME 4, the latter having been reported on in the preceding issue of the present report series, were performed simultaneously. The teaching phase of this study took place in April and May, 1970, and consisted of a series of six lessons in which a particular grammatical structure was taught, namely the passive voice. The pupils were in their fifth year of English (grade 8, approximately 15 years of age).

The independent variables of the experiment were three teaching methods, namely

Im	The Implicit method
Ee	The Explicit-English method
Es	The Explicit-Swedish method

Although the names of the teaching strategies are the same as in the previous studies (GUME 1-3) the teaching procedures were altered to some extent. As was also the case in GUME 4, the time for explanations in the present study varied between Ee and Es. A strong need was felt for the E methods to contain "optimal" explanations even if this meant a certain variation in explanation time, causing some looseness in experimental control. The Implicit method corresponds to an audio-lingual method without generalizations, the Explicit-English method corresponds to an audio-lingual method with generalizations in the target language, the Explicit-Swedish method corresponds to an audio-lingual method with explanations or generalizations in the source language; comparisons with corresponding structures in Swedish were also made.

In the study twenty-four school classes took part, twelve of which represented the more advanced course in English, sk (= särskild kurs),

and twelve the easier course, ak (= allmän kurs). Within each course four classes were taught according to each method /Im/Ee/Es/. The school classes were randomly assigned to teaching method, the only restriction on the procedure being that no two classes within the same course and school were allowed to get the same method.

Three parallel lesson series (Im/Ee/Es) were constructed, each consisting of six lessons. In order to control the teacher factor, "canned" lessons were used throughout the experiment. However, the teachers were, in a strictly prescribed way, instructed to take an active part in the work, especially in the case of oral drills; this was done by way of pointing, gestures, etc.

In rough outline the experimental schedule was as follows: tests of scholastic aptitude and listening comprehension were administered, distribution of lesson materials to the schools, Pre-test, the lesson series (i.e. the experiment proper), Post-test, Pupil and Teacher Attitude tests.

Progress during the experiment was measured as the difference between the Post-test and the Pre-test scores. The Criterion test was constructed so as to correspond to the particular objectives of the present investigation. It covered the particular grammatical structure taught and contained 94 items in all.

The test of scholastic aptitude was the so-called DBA-test (Differentiell BegåvningsAnalys = Differential Intelligence Analysis). The reason for administering this test was partly to use it as a background variable in some of the analyses and partly to divide the pupil population into three levels of ability and investigate interaction between teaching method and ability level.

In the statistical treatment of all data the two courses, sk and ak, were kept apart. Only pupils who had been present during at least five out of the six lessons were included; this is equal to stating that those who were absent from two or more lessons were not included in the calculations. Various checks on the drop-outs thus defined (absent two lessons or more) showed that they did not deviate from the experimental population in background variables; thus there is reason to believe that absence was due to chance (illness, visits to the school dentist, and the like). One significant difference was found:

in the advanced course the experimental population scored higher than the drop-outs, which may be taken as an indication that in the sk group the instructional program worked well - it was worthwhile being present during the lessons.

The standing of the experimental groups - sk and ak - on some relevant background variables (DBA, Grades, the National test in English) was checked. In the case of DBA and Grades the total group (sk + ak) are very close to the norm, in the case of the National test in English the sk group is somewhat above and ak about as much below the respective norms. On all measures the mean of the sk group is significantly above the mean of the ak group. With respect to social class the experimental population is biassed in so far as it contains disproportionately many pupils from social class 1.

The total progress in raw scores during the experiment was 7.88 points in sk and 3.72 in ak.

In a number of analyses of covariance Progress was the dependent variable and various background measures (DBA, PACT, Pre-test, the Standardized test in English) were used as covariates. Likewise, an analysis of covariance was performed with the Post-test as the dependent variable and the Pre-test as the covariate.

In the advanced course the three teaching methods, Im/Ee/Es, proved to be equally effective; the F-ratios were so low as to make consideration of tendencies among the absolute figures meaningless. In the easier course the Explicit-Swedish method was significantly superior to the two others in a number of analyses. However, the method differences in favour of Es should be interpreted with the utmost care for various reasons: The progress score in ak was grossly unreliable, the progress in general was limited, the superiority of Es is mainly found in two part tests where test effects, rather than differential progress, explain the Es superiority.

In order to investigate if there was any interaction between teaching method and ability level, the total population (sk + ak) was divided into three ability levels according to the scores on the DBA test (upper, middle, lower). An analysis of variance (two-way) was performed; no interaction was found. On the contrary a significant column (treatment)effect appeared (Es>).

The analyses mentioned thus far were made with individual scores as the unit of analysis. A complementary analysis was performed in ak with the school class mean as the unit of analysis. The difference in favour of the Es method was found at the school class level, too.

Two additional measures of Progress were calculated, both relating the pupil´s Progress score to his score on the Pre-test. The results obtained earlier with the raw Progress scores were duplicated; no differences in sk were found, but Es superiority in ak.

The pupils´ attitude to the project was fairly neutral in the advanced course; they considered the experimental teaching to be about the same as ordinary teaching - which remains to be defined. The ak pupils leaned towards the negative, the commonest response being that the lessons were rather boring and tedious. This attitude on the part of the ak group is natural since the lesson materials were identical for sk and ak, an unfortunate compromise striking the easier course the harder. However, for experimental reasons, the same teaching materials had to be accepted for the two courses.

The results in sk thus coincide completely with those found earlier in the GUME project. The results in ak favouring one particular method should be treated with caution for reasons given above. Viewed in the perspective of the intense debate in Sweden on foreign language teaching methodology, it is interesting to note that dramatic differences in favour of one teaching method still await detection.

BIBLIOGRAPHY

Earlier GUME Reports

Lindblad Torsten (1969) *Implicit and Explicit. An Experiment in Applied Psycholinguistics, Assessing Different Methods of Teaching Grammatical Structures in English as a Foreign Language.* GUME-projektet 1. Lärarhögskolan och Universitetet i Göteborg.

Carlsson Ingvar (1969) *Implicit and Explicit. An Experiment...* GUME-projektet 2.

Olsson Margareta (1969) *Implicit and Explicit. An Experiment...* GUME-projektet 3.

Levin Lennart (1969) *Implicit and Explicit. A Synopsis of Three Parallel Experiments in Applied Psycholinguistics, Assessing Different Methods of Teaching Grammatical Structures in English as a Foreign Language.* The GUME Project. Department of Educational Research, Gothenburg School of Education.

Lindblad & Levin (1970) *Teaching Grammar. An Experiment in Applied Psycholinguistics, Assessing Three Different Methods of Teaching Grammatical Structures in English as a Foreign Language.* The GUME Project. Department of Educational Research, Gothenburg School of Education.

Other References.

Åhlström K (1894) *Engelsk Språklära.* Stockholm: Fritze.

Aleamoni L M & Spencer R E (1969) "An Evaluation of the Pennsylvania Foreign Language Project". *The Modern Language Journal.* Vol. LIII, No. 6 (October, 1969).

Anastasi A (1958) *Differential Psychology.* New York: Macmillan.

Anderson R C (1969) "The Comparative Field Experiment: An Illustration from High School Biology". *Proceedings of the 1968 Invitational Conference on Testing Problems,* Educational Testing Service.

Andrakammarvalet i Göteborg 1968. Stickprovsundersökning. (A Sample Study of the Gothenburg Electorate in The Parliamentary Election in 1968) Göteborgs stads statistiska kontor. U 1969:2.

Anvisningar till standardproven för grundskolans årskurs 8. (1969) Stockholm: Skolöverstyrelsen. (Manual for the Standardized Tests in Grade 8).

Ausubel D (1964) "Adults Versus Children in Second Language Learning: Psychological Considerations". *MLJ*, Vol. XLVIII, pp. 420-4.

Betänkande afgifvet den 8 December 1902 af den för utredning af vissa frågor rörande de allmänna läroverken den 26 Maj 1899 i nåder tillsatta kommittē. (Kommittē-Betänkandet). Stockholm.

Bloomfield L (1933) *Language and Language Teaching*. New York: Harcourt, Brace & World.

Bosco J F and Di Pietro R J (1970) "Instructional Strategies: Their Psychological and Linguistic Bases". *IRAL*, Vol. VIII, No. 1, pp. 1-19.

Brooks N (1960) *Language and Language Learning*. New York: Harcourt, Brace & World.

Campbell D T & Stanley J C (1963) "Experimental and Quasi-experimental Designs for Research on Teaching". In *Handbook of Research on Teaching*, ed. Gage (1963), pp. 171-246.

Campbell R N (1970) "An Evaluation and Comparison of Present Methods for Teaching English Grammar to Speakers of Other Languages". *TESOL*, Vol. 4, No. 1, pp. 37-48.

Carroll J B (1965) "The Contributions of Psychological Theory and Educational Research to the Teaching of Foreign Languages". *MLJ*, Vol. XLIX, No. 5, pp. 273-81.

Carroll J B (1969) "What Does the Pennsylvania Foreign Language Research Project Tell Us?" *Foreign Language Annals*, Vol. 3, No. 2 (December, 1969).

Center for Curriculum Development (1970) *"Stages in Language Learning"*. Philadelphia: CCD. Work paper.

Centre for Information on Language Teaching (CILT), Reports and Papers 2:(1969). *Aims and Techniques: Language-teaching Methods and Their Comparative Assessment.* London: CILT.

Chastain K (1969) "The Audio-Lingual Habit Theory Versus the Cognitive Code-Learning Theory: Some Theoretical Considerations". *IRAL*, Vol. VII, No. 2, pp. 97-106.

Cronbach L (1963) "Evaluation for Course Improvement". *Teachers' College Record*, 64, No. 8 (May, 1963).

Dahlgren O (1947) *Bidrag till språkfaktorernas psykologi.* Göteborg: University of Gothenburg.

Dunlop I (1970) *Practical Techniques in the Teaching of Oral English.* Stockholm: Almqvist & Wiksell.

Edwardsson R (1970) *Språkdebatten 1969-1970.* Stockholm: Utbildningsförlaget.

Elfstrand D (1897) *An English Grammar for the Use of Swedish Schools.* Stockholm: Bille.

Fries C C (1946) *Teaching and Learning English as a Foreign Language.* Ann Arbor: University of Michigan Press.

Gage N L (Ed) (1963) *Handbook of Research on Teaching.* Chicago: Rand McNally & Co.

Grittner F M (1968) "Letter to the Editor". *Newsletter of the National Association of Language Laboratory Directors*, Vol. III, No. 2 (December, 1968).

Hocking E (1969) "The Laboratory in Perspective: Teachers, Strategies, Outcomes". *MLJ*, Vol. LIII, No. 6 (October, 1969).

Härnqvist K (1969 a) *Nya normer för DBA årskurs 6.* Stockholm: Skandinaviska Testförlaget. (New Norms for DBA, grade 6).

Härnqvist K (1969 b) *Resultat från undersökningar med DBA och DBA-liknande test.* Stockholm: Skandinaviska Testförlaget. (Results from Investigations with DBA and Similar Tests).

Jacobovits L (1968) "Implications of Recent Psycholinguistic Developments for the Teaching of a Second Language". *Language Learning*, Vol. XVIII, Nos 1 and 2, pp. 89-109.

Jespersen O (1961) *How to Teach a Foreign Language*. London: George Allen & Unwin Ltd. (First published in 1904).

Lado R (1961) *Language Testing*. London: Longmans.

Larsson L & Sandgren B (1968) *En studie av kreativitetsutvecklingen inom årskurserna 4-9 samt en undersökning av kreativitetens samvariation med intelligens*. (A Study of the Development of Creativity within Grades 4-9 and an Investigation of the Covariation of Creativity with Intelligence). Unpublished.

Läroplan för grundskolan (1962) *"Lgr 62"* (Authorized Curriculum for Swedish Schools) Stockholm: Skolöverstyrelsen.

Läroplan för grundskolan (1969) *"Lgr 69"* (ditto) Stockholm: Skolöverstyrelsen.

Läroplan för grundskolan, *Lgr 69, Supplement Engelska* (1969). (Ditto, Supplementary Comments on English as a Foreign Language) Stockholm: Skolöverstyrelsen.

Läroplan för gymnasiet (1965) (Authorized Curriculum for the Top Forms in Sweden). Stockholm: Skolöverstyrelsen.

Lim K-B (1968) *Prompting Vs Confirmation, Pictures Vs Translations, and other Variables in Children's Learning of Grammar in a Second Language*. Princeton. Unpublished thesis.

Moulton W (1961) "Linguistics and Language Teaching in the United States 1940-1960", in *Trends in European and American Linguistics*, eds. Mohrmann, Sommerfelt, Whatnough. Utrecht: Spectrum Publishers.

Nachman M & Opochinsky S (1958) "The Effects of Different Teaching Methods: a Methodological Study". *J. educ. Psychol.*, 49 (1958), pp. 245-249.

Ney J W (1968) "The Oral Approach: A Re-Appraisal". *Language Learning*. Vol. XVIII, Nos 1 and 2, pp. 3-13.

Northeast Conference (1958) *The Language Teacher*. Ed. Levy. New York: MLA Materials Center.

Otter H S (1968) *A Functional Language Examination*. London: Oxford University Press.

Otto F (1969) "The Teacher in the Pennsylvania Project". *MLJ*, Vol. LIII, No. 6 (October, 1969).

Palmer H E (1965) *The Principles of Language-Study*. London: Oxford University Press. (First published in 1921).

Pattison B (1970) "Research Priorities". *English Language Teaching*, Vol. XXV, No. 1, pp. 3-10.

The Random House Dictionary of the English Language (1968) New York: Random House, Inc.

Rivers Wilga (1964) *The Psychologist and the Foreign-Language Teacher*. Chicago: The University of Chicago Press.

Rivers Wilga (1968) *Teaching Foreign-Language Skills*. Chicago: The University of Chicago Press.

Scriven M (1968) *The Methodology of Evaluation*. AERA Monograph Series of Curriculum Evaluation, No. 1, Chicago, 1968.

Smith P D & Berger E (1968) *An Assessment of Three Foreign Language Teaching Strategies*. Washington: US Department of Health, Education & Welfare.

Smith P D & Baranyi H A (1968) *A Comparison Study of the Effectiveness of the Traditional and Audiolingual Approaches to Foreign Language Instruction Utilizing Laboratory Equipment*, Washington: US Department of Health, Education & Welfare.

Smith P D (1969 a) *A Comparison Study of the Effectiveness of the Traditional and Audiolingual Approaches to Foreign Language Instruction Utilizing Laboratory Equipment*. Washington: US Department of Health, Education & Welfare.

Smith P D (1969 b) "The Pennsylvania Foreign Language Research Project: Teacher Proficiency and Class Achievement in Two Modern Languages". *Foreign Language Annals*, Vol. 3, No. 2 (December, 1969).

Smith P D (1969 c) *In Reply to the October, 1969, Modern Language Journal*. A talk by Dr. Philip D. Smith, Jr., 3rd Annual Meeting of the American Council on the Teaching of Foreign Languages, New Orleans, November 28, 1969. Mimeographed manuscript, a revised version of which is included in Smith, 1969 a.

Stanley J C (1970) "Controlled Experimentation: Why Seldom Used in Evaluation?" In *Proceedings of the 1969 Invitational Conference on Testing Problems*, Educational Testing Service, pp. 104-108.

Statistiska meddelanden U 1969:5 and U 1970:5. Stockholm: Statistiska Centralbyrån.

Stephens J M (1967) *The Process of Schooling*. New York: Holt, Rinehart and Winston.

Travers R M W (1960) *An Introduction to Educational Research*. London: Macmillan.

Valette Rebecca M (1967) *Modern Language Testing*. New York: Harcourt, Brace & World.

Wiley D E (1969) "A Methodological Review of the Pennsylvania Foreign Language Research Project". *Foreign Language Annals*, Vol. 3, No. 2 (December, 1969).

Wittrock M C & Wiley D E (Ed) (1970) *The Evaluation of Instruction*. New York: Holt, Rinehart & Winston. Forthcoming.

List of Appendices

Appendix A: The Criterion test given as pre- and post-test.

Appendix B: The Pupil Attitude Test administered at the end of the project immediately after the post-test.

Appendix C: The Teacher Attitude Test, consisting of two parts, one concerning general methodology, the other various aspects of the project, answered at the time of the pupil attitude test.

Appendix D: A list of participating teachers in alphabetical order.

Appendix E: Illustrations of the grammatical explanations used in the Explicit-English group.

Appendix F: Complementary statistics: Means and standard deviations for the experimental population (sk + ak, sk, ak) and per school class. Pre-, Post-test and Progress correlations for sk and ak.

Appendix A

THE CRITERION TEST

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

Namn

Klass

Skola

Eng. lärare

E n g e l s k a å k 8

The Passives

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

P r o v I

I följande stycke har en del ord utelämnats. En streck betyder ett utelämnat ord. När du nu läser igenom texten skall du sätta in det ord du tycker passar bra i meningen.

The Swan Lake by Tchaikovsky _____ written by a famous composer, but very often the music for ballets _____ composed by fairly unknown musicians, and many of them have _____ forgotten nowadays. In the many books that have _____ published on ballet the music _____ hardly mentioned at all. What music the famous dancers of the past danced to can only _____ discovered in the museums. People who _____ thrilled by the dancers from the Bolshoi Theatre in Moscow, for instance, seldom talk about how the orchestra played. In the programmes of the theatres you will find that the names of the dancers _____ printed (print = trycka) with much bigger letters than the name of the composer. When the famous dancer Nureyev was in Stockholm he _____ received with much more publicity than any composer could hope for. And at the same time it is very probable that the name of Tchaikovsky will _____ remembered much longer than Nureyev´s. It would be more natural if people _____ as much attracted by the music as by the dancers.

VÄND EJ BLAD FÖRRÄN DU BLIR TILLSAGD!

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

P r o v 2

I stället för att säga:
The girl has been found by the police
kan man säga:
The police have found the girl.

Ändra nu följande meningar på samma sätt. I den undre meningen får du några ord till hjälp och sedan fyller du i resten själv. Det finns streck för varje ord du ska sätta in. Nu börjar vi:

1. The diamonds have already been found by the thief.
   The thief _____ already _____ the diamonds.

2. The desk has already been painted by the children.
   The children _____ already _____ the desk.

3. The letters have already been written by grandmother.
   Grandmother _____ already _____ the letters.

4. The chair has already been sold by the parents.
   The parents _____ already _____ the chair.

5. The film has already been forgotten by the children.
   The children _____ already _____ the film.

6. The food has already been cooked by mother.
   Mother _____ already _____ the food.

7. The elephant has already been shot by the hunters.
   The hunters _____ already _____ the elephant.

(skrivningen fortsätter på nästa sida)

8. The washing-up has already been done by the boys.

   The boys _____ already _____ the washing-up.

9. The car has already been bought by Father.

   Father _____ already _____ the car.

10. The astronaut has already been seen by millions of people.

    Millions of people _____ already _____ the astronaut.

VÄND INTE BLAD FÖRRÄN DU BLIR TILLSAGD!

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

Prov 3

DOCTOR DOLITTLE AND HIS FRIENDS

1. a) The Doctor gave the cows milk in payment.
   b) The Doctor didn´t give the cows milk in payment.
   c) The cows gave the Doctor milk in payment.
   d) The cows didn´t give the Doctor milk in payment.

2. a) Bellowes hadn´t killed her husband.
   b) Bellowes had killed her husband.
   c) Her husband had killed Bellowes.
   d) Her husband hadn´t killed Bellowes.

3. a) The foxes protected the skunks.
   b) The foxes didn´t protect the skunks.
   c) The skunks protected the foxes.
   d) The skunks didn´t protect the foxes.

4. a) A very unusual animal led out the Doctor.
   b) The Doctor didn´t lead out the unusual animal.
   c) The unusual animal didn´t lead out the Doctor.
   d) The Doctor led out the unusual animal.

5. a) It had sent one of the Doctor´s friends.
   b) It hadn´t sent one of the Doctor´s friends.
   c) One of the Doctor´s friends had sent it.
   d) One of the Doctor´s friends hadn´t sent it.

6. a) Dr Dolittle didn´t take the pushmi-pullyou to a circus.
   b) Dr Dolittle took the pushmi-pullyou to a circus.
   c) The pushmi-pullyou took Dr Dolittle to a circus.
   d) The pushmi-pullyou didn´t take Dr Dolittle to a circus.

7. a) Her husband had caught the sailors.
   b) The sailors had caught her husband.
   c) Her husband hadn´t caught the sailors.
   d) The sailors hadn´t caught her husband.

8. a) Dr Dolittle didn´t carry Sophie.
   b) Dr Dolittle carried Sophie.
   c) Sophie carried Dr Dolittle.
   d) Sophie didn´t carry Dr Dolittle.

9. a) They thought a woman had murdered Dr Dolittle.
   b) They thought a woman hadn´t murdered Dr Dolittle.
   c) They thought Dr Dolittle hadn´t murdered a woman.
   d) They thought Dr Dolittle had murdered a woman.

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

P r o v 4

I följande meningar skall antingen "will be" eller "will have" sättas in. I högra kanten av detta papper hittar du dessa ord. Sätt ett kryss i kolumnen under det uttryck du tycker passar bäst i meningen i fråga. Denna skrivning består av två papper. Du vänder alltså blad när du är färdig med första sidan.

	will have	will be
1. He thinks he ... chosen President next year.	1 a)	b)
2. The students ... passed their exams before the end of June.	2 a)	b)
3. The job ... finished before Christmas.	3 a)	b)
4. The prize ... won by the best pupil.	4 a)	b)
5. The money ... put in the bank.	5 a)	b)
6. The children ... told a bedtime story.	6 a)	b)
7. The acrobat ... trained the monkeys perfectly in a year.	7 a)	b)
8. The building ... admired when it is finished.	8 a)	b)
9. The soldiers ... shown their new caps to the general before next Sunday.	9 a)	b)
10. The fire ... destroyed the town before the fire brigade arrives.	10 a)	b)
11. The house ... painted green.	11 a)	b)
12. The pictures ... hung in the palace not later than Friday.	12 a)	b)
13. The children ... done their homework before eight.	13 a)	b)
14. The gardener ... watered the flowers by the time I have cut the lawn.	14 a)	b)
15. A hot bath ... prepared for you when you come back from your long railway trip.	15 a)	b)
16. At Operakällaren a wonderful dinner ... cooked by Tore Wretman when Grandmother has her eightieth birthday.	16 a)	b)
17. The gates ... shut every evening at ten o´clock.	17 a)	b)

(Skrivningen fortsätter på nästa sida)

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

	will have	will be
18. She ... given a silver spoon on her birthday.	18 a)	b)
19. In this small house a new book ... written every year.	19 a)	b)
20. The story ... repeated at every party.	20 a)	b)
21. By next Sunday the young man ... told the girl that he loves her.	21 a)	b)
22. The contract ... signed next week.	22 a)	b)
23. The hotel ... paid as quickly as possible for the two rooms.	23 a)	b)
24. The captain ... sold the ship before we can get enough money to buy it.	24 a)	b)
25. The professors ... warned people against swimming in this lake before it is summer.	25 a)	b)
26. Too much money ... spent on this house before people can move into it.	26 a)	b)
27. By the time the child is asleep, the mother ... sent for the doctor.	27 a)	b)
28. If this question it put to them, the pupils ... answered it within a second.	28 a)	b)
29. The troops ... called out very soon.	29 a)	b)
30. The police ... found the pearls and the diamonds before the thieves have time to sell them.	30 a)	b)
31. Mini-skirts ... forgotten when long skirts are fashionable (modern) again.	31 a)	b)
32. I think that the children ... noticed that the road is very bad.	32 a)	b)
33. If he is not careful in the jungle, the lions ... killed him before he has time to take out his gun.	33 a)	b)
34. Kind people ... helped the poor family by next month.	34 a)	b)
35. The rich lady ... given away her money before she is old.	35 a)	b)
36. The President ... spoken by the end of next month.	36 a)	b)
37. All the tourists ... returned to their country by the time the winter sets in.	37 a)	b)
38. People ... picked up all the money I lost on my way to school.	38 a)	b)
39. When you leave school, you ... studied English for many years.	39 a)	b)
40. The captain ... inspected every corner before the holidays are over.	40 a)	b)

155 S L U T

VÄND INTE BLAD FÖRRÄN DU BLIR TILLSAGD!

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

P r o v 5

1. What is your name?

   I _____ _____ (call) Piggy by my friends.

2. Did you walk home yesterday?

   No, I _____ _____ (take) home by my cousin.

3. Do you know where the cups are?

   Yes, they _____ always _____ (put) in the cupboard.

4. Where did you spend your holiday?

   It _____ _____ (spend) with my family on Bornholm.

5. Can you study foreign languages in Swedish schools?

   Yes, English, German, and French _____ _____ (teach) in Swedish schools.

6. Did the children take any apples?

   No, all the apples _____ _____ (steal) by the thieves.

7. Had you enough money for the bicycle?

   No, it _____ _____ (pay) for by my father.

8. Did you hear a noise in your room last night?

   No, but all sorts of noises _____ _____ (hear) in this house at night.

9. Does anybody visit that old museum?

   Yes, it _____ _____ (visit) by many people on Sundays.

10. Shouldn´t we repair the house next summer?

    But it _____ _____ (repair) last summer.

(Skrivningen fortsätter på nästa sida)

11. Do the French speak German?

    The French speak French. German _____ _____ (speak)

    in Germany.

12. What was the girl doing outside the school?

    She was waiting. The doors _____ normally _____ (open)

    by the caretaker, but today he was ill.

13. Do you know when Selma Lagerlöf wrote Gösta Berlings saga?

    Gösta Berlings saga ____ _____ (write) in 1891.

14. Didn´t you see the boy?

    I didn´t see the boy, but he ____ _____ (see) by all the others.

15. What are you going to wear at the party?

    I want a new dress, but I don´t know where the best ones

    _____ _____ (sell).

17. Is this a new house?

    No, it _____ _____ (build) in 1851.

S L U T

VÄND INTE BLAD FÖRRÄN DU BLIR TILLSAGD!

GUME-projektet 5
Lärarhögskolan i Göteborg
Margareta Olsson

P r o v 6

För att uttrycka att kaffe serveras klockan åtta kan man på engelska använda två fraser, nämligen:

a) They serve coffee at eight

eller

b) Coffee is served at eight.

Fyll nu i följande meningar enligt samma mönster.

1a) They eat plumpudding at Christmas in England.

1b) Plumpudding .....

2a) They play cricket in summer in England.

2b) ....

3a) They make matches at Jönköping.

3b) .....

4a) They respect aristocrats in England.

4b) .....

5a) They sell beautiful clothes in Paris.

5b) .....

6a) They export fine cameras from Japan.

6b) .....

7a) They drink vodka in Russia.

7b) .....

8a) They speak Spanish in Spain.

8b) .....

(Skrivningen fortsätter på nästa sida)

9a) They teach swimming in the schools in Sweden.

9b) .....

10a) They study French in the evenings.

10b) .....

11a) They visited their Grandmother very often.

11b) .....

S L U T

GUME-projektet
GUME 5 1970
Lärarhögskolan i Göteborg
Margareta Olsson

Prov 3

DOCTOR DOLITTLE AND HIS FRIENDS

This is a story taken from one of the books about Dr Dolittle. The Doctorlivedin a small place in England. To start with he was a Doctor for people, but then he became a Doctor for animals. Here is what happened one morning:

1. By 5. 30 in the morning the front hall was filled with patients that mooed, bleated, and neighed (som råmade, bräkte och gnäggade) outside the library door. All of them had come without people. They didn't pay the Doctor in pounds and shillings as people do. TheDoctor was given milk by the cows in payment for the medicine.

2. Tommy helped Polynesia, the parrot, by calling out, "Who's next?" and led the patients in and out of the library. A fox named Sheila brought in her three little ones. Sheila was unhappy because her husband had been killed by Bellowes.

3. When Doctor Dolittle was examining a horse, Bellowes arrived and said that the Doctor had stolen his horse. Then he saw the fox family and cried to his dogs to pursue them. The Doctor just laughed because he had imported some skunks, and as they smelt so strong and so bad, the foxes were protected by the skunks.

4. The dogs couldn't catch any foxes and Bellowes went home again very angry. Everybody was still laughing at the fox and skunk episode when a big box arrived for the Doctor. One end of the box was opened and a very unusual animal was led out by the Doctor.

5. The animal, which had two heads, was called a pushmi-pullyou. Each head looked in an opposite direction. It had been sent by one of the Doctor's friends.

6. If Doctor Dolittle showed the animal for money he could earn enough to set out on a long journey. Because of that the pushmi-pullyou was taken to a circus by Dr Dolittle.

7. At the circus Dr Dolittle got to know a beautiful seal called Sophie (en vacker säl, som hette Sofi). She was very unhappy because when she was caught by the sailors her husband had not been caught too.

8. The Doctor decided to help Sophie to get to the sea. One dark night they ran away from the circus. Before they found a train to go by Sophie was carried by the Doctor for several hous.

9. When they got to the sea, the Doctor flung Sophie into the water. However, when he turned to go back, he saw that two policemen were watching him, and they suddenly took hold of his arms. They thought that a woman had been murdered by Dr Dolittle.

+ + +

Appendix B

THE PUPIL ATTITUDE TEST

Lärarhögskolan i Göteborg
GUME-projektet, GUME 5
Margareta Olsson 1970

Intresse för olika skolämnen

Namn: _____ Klass: _____

Skola: _____

Engelsklärare: _____

Jag läser ____ allmän ____ särskild kurs i engelska.

==========

Sätt ett kryss (x) för varje ämne inom parentesen under den pil som bäst visar hur du tycker om det ämnet! Tänk efter inte bara hur du tycker just idag utan hur du brukar tycka.

Hoppa inte över något ämne som du har!

	Nästan alltid roligt ↓	Mera roligt än tråkigt ↓	Mera tråkigt än roligt ↓	Nästan alltid tråkigt ↓
Svenska	( )	( )	( )	( )
Matematik	( )	( )	( )	( )
Engelska	( )	( )	( )	( )
Kristendomskunskap	( )	( )	( )	( )
Samhällskunskap	( )	( )	( )	( )
Biologi	( )	( )	( )	( )
Fysik	( )	( )	( )	( )
Musik	( )	( )	( )	( )
Teckning	( )	( )	( )	( )
Slöjd	( )	( )	( )	( )
Hemkunskap	( )	( )	( )	( )
Gymnastik	( )	( )	( )	( )
Tyska	( )	( )	( )	( )
Franska	( )	( )	( )	( )
Maskinskrivning	( )	( )	( )	( )

Lärarhögskolan i Göteborg
GUME-projektet GUME 5
MO 1970

Du har under den senaste tiden varit med i det s k GUME-projektet, vilket har inneburit att du dels fått ett antal olika prov och dels fått följa sex lektioner på bandspelare. Vi vill nu höra litet om vad du tyckt om det här. Svara på alla frågorna; svara med kryss (x) eller korta meningar.

==========

1 Jag har varit med på _____ av de sex lektionerna.

2 Det som var bra med GUME-lektionerna var att

_____

_____

_____

3 Det som inte var bra med GUME-lektionerna var att

_____

_____

_____

4 På de här timmarna lärde jag mig engelska

_____ mycket bättre än på vanliga timmar

_____ något bättre än på vanliga timmar

_____ ungefär som på vanliga timmar

_____ något sämre än på vanliga timmar

_____ mycket sämre än på vanliga timmar

5 De här timmarna var

_____ mycket roligare än vanliga timmar

_____ något roligare än vanliga timmar

_____ ungefär som vanliga timmar

_____ något tråkigare än vanliga timmar

_____ mycket tråkigare än vanliga timmar

6. Det som var roligare var att _____

_____

7 Det som var tråkigare var att _____

_____

8 Tiden under de här timmarna verkade gå

_____ mycket fortare än under vanliga timmar

_____ något fortare än under vanliga timmar

_____ ungefär som under vanliga timmar

_____ något långsammare än under vanliga timmar

_____ mycket långsammare än under vanliga timmar

9 Efter de här timmarna kände jag mig

_____ mycket tröttare än efter vanliga timmar

_____ något tröttare än efter vanliga timmar

_____ ungefär som efter vanliga timmar

_____ något mindre trött än efter vanliga timmar

_____ mycket mindre trött än efter vanliga timmar

10 (Om du var trött:) Det som gjorde mig trött var:

_____

_____

11 Jag tyckte att ljudet i allmänhet var

_____ mycket bra och lätt att höra

_____ bra

_____ rätt dåligt

_____ mycket dåligt och svårt att höra

12. (Denna fråga skall du bara besvara om du hade gröna eller röda blad med förklaringar på i dina buntar på lektionerna)

De förklaringar vi fick tyckte jag

_____ gjorde det mycket lättare att förstå

_____ gjorde det något lättare att förstå

_____ inte gjorde någon skillnad

_____ gjorde det något svårare att förstå

_____ gjordet det mycket svårare att förstå

13 De muntliga övningarna, då vi skulle prata själva, tyckte jag var

_____ mycket bra
_____ bra
_____ rätt dåliga
_____ mycket dåliga

} därför att _____

14 De skriftliga övningarna tyckte jag var

_____ mycket bra
_____ bra
_____ rätt dåliga
_____ mycket dåliga

} därför att _____

15 Lästexterna tyckte jag var

_____ mycket bra
_____ bra
_____ rätt dåliga
_____ mycket dåliga

} därför att _____

16 Ytterligare kommentarer som jag skulle vilja framföra:

_____

_____

_____

_____

Appendix C

THE TEACHER ATTITUDE TEST

Lärarhögskolan i Göteborg
GUME-projektet GUME 5
MO 1970

Lärarenkät I - allmän metodik

Detta formulär kan ifyllas anonymt om Du så önskar och insändas separat. Skriv svaren på maskin om Du vill. Vi kommer givetvis inte att försöka spåra någon men tycker det vore intressant att kartlägga lärarinställningen till de angivna frågorna. Vi hoppas på hundraprocentig svarsfrekvens.

Namn (ej obligatoriskt): _____

1 man / kvinna 2 Ålder: 21-30 31-40 41-50 51-60 61-

2 Utbildning: ___ folkskollärare
___ vidareutbildning i engelska
___ fil mag
___ annan utbildning (Vad? _____)

4 Antal betyg i engelska ___

5 Erfarenhet: ___ mellanstadiet i engelska (_____ antal år)
___ högstadiet (_____ år) (Tre år med tjänst på både högst och gymn markeras som 3+3 år)
___ gymnasium (+fackskola) (_____ år)
___ annat (Vad? _____, _____ år)

Ämnen: betyg i _____ _____ _____ _____ _____
erfarenhet av undervisning _____ _____ _____ _____ _____
ty fr nord språk sv ry annat

6 Praktisk lärarutbildning år _____ (folkskollärarex, provår, lärarhögskola)

==========

De tre metoderna, som ingår i försöket är Im (ren strukturövning), Ee (förklaringar på engelska) och Es (förklaringar på svenska).

7 Jag tror - utan att i detalj känna till de andra två metoderna - att vi kommer att lyckas bäst med

		Im	Ee	Es
bland svaga	___			
bland medel	___			
bland duktiga	___			

8 Jag brukar nog själv i åk 8 följa vad som närmast torde motsvara

Im Ee Es annat (Vad? _____)

9 Jag tycker att eleverna på högstadiet bör ha en grammatiklärobok. Ja / nej

10 Jag tycker att man bör ge grammatiska förklaringar: (stryk under)

varje lektion rätt ofta och regelbundet någon gång aldrig

11 Om grammatisk förklaring skall ges, så bör den ges:

_____ a) på svenska

_____ b) på engelska

12 Om grammatiska förklaringar används bör de ges

_____ a) av läraren, snabbt och koncist

_____ b) av någon elev och rundas av efteråt av läraren

13 Jag anser att undervisningen bör föras till ca _____% på engelska i allmän kurs

14 Jag anser att undervisningen bör föras till ca _____% på engelska i särskild kurs

Lärarhögskolan i Göteborg
GUME-projektet GUME 5
MO 1970

Lärarenkät II - synpunkter på projektet

Vi ber Dig fylla i detta formulär så omsorgsfullt och noga som möjligt. Använd gärna baksidan eller extrablad för att ge fylliga kommentarer.

Nam: _____ Skola: _____

Jag har en _____ kurs som undervisades efter Im/Ee/Es - metoden.

1 Bra med den metodik som min klass undervisades efter var (om Du hade två klasser med så dela upp synpunkterna):

_____

_____

_____

_____

_____

2 Mindre bra eller dåligt var: (jfr frågorna nedan innan Du svarar)

_____

_____

_____

_____

3 Ange nedan kortfattat. Din åsikt om:

a) De grammatiska förklaringarna (för E-grupperna)

_____

_____

_____

b) De muntliga övningarna

_____

_____

_____

c) De skriftliga övningarna

_____

_____

_____

d) Läsövningarna

4 Om tempot i lektionerna - pauslängder och talhastighet - anser jag:

5 Om den tekniska kvaliteten på inspelningarna anser jag:

6 Elevernas reaktion jämfört med vanlig undervisning synes vara beträffande

a) intresse:

b) disciplin:

c) inlärningseffekter:

7 Om för- och efterprovet anser jag:

8 Kommentarer - positiva och negativa - till de enskilda lektionerna (gärna lektionsvis för alla sex):

_____

_____

_____

_____

_____

_____

_____

9 På det hela taget tycker jag att tiden som experimentet tagit varit:

_____ i det närmaste helt bortkastad

_____ tämligen outnyttjad

_____ ungefär som vanligt

_____ tämligen effektivt utnyttjad

_____ mycket väl utnyttjad

10 Den metod som mina elever fått pröva anser jag vara:

_____ dödfödd

_____ användbar i framtiden med följande ändringar: _____

_____

_____

11 Ytterligare kommentarer:

_____

_____

_____

_____

_____

_____

_____

Ed/ 1970. 04. 16

Appendix D

PARTICIPATING TEACHERS

List of Participating Teachers (in Alphabetical Order)

sk:

Name of teacher	School	
Ulla Arnholm	Utmarksskolan	Göteborg
Ann-Marie Blom	Kvarnbyskolan	Mölndal
Uno Bredinge	Ånässkolan	Göteborg
Lars Dahllöf	Böskolan	Göteborg
Ingrid Elmén	Kärralundsskolan	Göteborg
Lilian Falkenland	Nya Lundenskolan	Göteborg
Ralph Fredriksson	Sannaskolan	Göteborg
Inger Friberg	Kärralundsskolan	Göteborg
Sigrid Groot	Kvarnbyskolan	Mölndal
Karin Lundborg	Nya Lundenskolan	Göteborg
Harriet Lundvall	Åbyskolan	Mölndal
Lilian Zachrisson	Kvarnbyskolan	Mölndal

ak:

Karl Anders Augustsson	Centralskolan	Stenungssund
Roy Ernered	Utmarksskolan	Göteborg
Inger Friberg	Kärralundsskolan	Göteborg
Kjell Gisslin	Utmarksskolan	Göteborg
Sven Hallbert	Sannaskolan	Göteborg
Anita Hellberg	Bleketskolan	Tjörn
Åke Hult	Utmarksskolan	Göteborg
Eva Karlsson	Flatåsskolan	V. Frölunda
Bertil Liljedahl	Ekebäcksskolan	V. Frölunda
Inger Nyström	Sannaskolan	Göteborg
Doht Persson	Flatåsskolan	V. Frölunda
Anna-Lisa Svensson	Gärdsåsskolan	Göteborg

Appendix E

ILLUSTRATIONS OF THE GRAMMATICAL EXPLANATIONS USED IN THE EXPLICIT-ENGLISH GROUP

III:Ee

A

Active: Barbra Streisand wore this necklace yesterday.
1 2 3
(huvudverb: wear)

Passive: This necklace was worn by Barbra Streisand yesterday.
3 2 1
(hjälpverb: was) (de tre formerna på huvudverbet: wear, wore, worn)

B

Active: Barbra Streisand has worn this necklace all the week.
1 2 3
(hjälpverb: has) (huvudverb: wear)

Passive: This necklace has been worn by all the week
3 2 1
(hjälpverb: has been) (huvudverb: wear)

C

Aktive: Mrs Barker has cleaned the shoes.
1 2 3
(hjälpverb: has) (huvudverb: clean)

Passive: by .
3 2 1

A
Active: They (1) have watered (2) the flowers. (3)
Passive: The flowers (3) have been watered (2) (1).

B
Active: (1) have posted (2) (3).
Passive: The letter (3) has been posted (2) (1).

C
Active: (1) have cleaned (2) (3)
Passive: The shoes (3) have been cleaned (2) (1).

A

Active: They (1) will tell (2) the other doctors (3) about it.

Passive: The other doctors (3) will be told (2) about it. (1)

B

Active: They (1) will not see (2) Benskin (3) with the night nurse.

Passive: Benskin (3) will not be seen (2) with the night nurse. (1)

C

Active: (1) will forget (2) everything (3) before next week.

Passive: Everything (3) will be forgotten (2) before next week. (1)

Appendix F

DESCRIPTIVE STATISTICS FOR THE EXPERIMENTAL POPULATION
(sk + ak, sk, ak)

DESCRIPTIVE STATISTICS PER SCHOOL CLASS; sk N = 12

DESCRIPTIVE STATISTICS PER SCHOOL CLASS; ak N = 12

PRE-, POST-TEST AND PROGRESS CORRELATIONS; sk N = 235

PRE-, POST-TEST AND PROGRESS CORRELATIONS; ak N = 152

Means and Standard Deviations for the Experimental Population (sk + ak, sk, ak).

	sk + ak			sk			ak		
	N	$\bar{x}$	s	N	$\bar{x}$	s	N	$\bar{x}$	s
DBA Verbal	334	5.08	1.83	214	5.87	1.63	120	3.67	1.19
DBA Inductive	334	5.14	1.99	214	5.89	1.73	120	3.81	1.73
DBA Spatial	334	5.21	2.08	214	5.59	2.08	120	4.54	1.89
DBA Total	334	50.98	10.22	214	55.24	8.80	120	43.37	7.91
Grades English	381	3.15	0.97	233	3.33	1.00	148	2.86	0.86
Grades Swedish	381	3.02	0.95	233	3.47	0.84	148	2.30	0.60
Grades Maths	381	3.00	1.03	233	3.21	1.06	148	2.66	0.89
Grades Total	381	27.42	7.45	233	29.95	7.42	148	23.43	5.50
Std. Test EL	377	15.62	6.53	229	18.88	5.57	148	10.56	4.32
Std. Test EM	376	15.40	6.48	228	19.12	5.31	148	9.68	3.03
Std. Test EA	377	14.30	5.64	229	11.30	3.83	148	18.95	4.79
Std. Test Total	377	45.35	13.05	228	49.41	12.81	148	39.18	10.49
PACT	366	47.54	5.88	227	50.19	3.10	139	43.19	6.72
Pre-test	387	48.32	18.34	235	59.11	14.83	152	31.52	7.08
Post-test	387	54.48	19.79	235	66.99	14.13	152	35.24	8.45
Progress	387	6.25	7.84	235	7.88	8.04	152	3.72	6.80
Pupil Attit.	340	22.06	4.44	200	22.55	4.43	140	21.34	4.38
Grades German	203	2.99	0.95	159	3.09	0.93	44	2.64	0.94
Grades French	69	3.49	1.15	68	3.51	1.14	1		

School Class Means in Certain Variables. sk; N = 12.

Method	School Class No.	N	DBA	Grades	Std Test	PACT	Pre-test	Post-test	Pro-gress	Pupil Attitude
Im	01	18	57.67	31.00	52.89	50.69	63.28	71.00	7.72	23.11
Im	02	19	58.21	32.37	59.00	50.37	63.68	70.79	7.11	22.33
Im	03	18	59.73	28.41	46.78	50.71	63.39	72.28	8.89	26.47
Im	04	15	52.54	28.50	39.40	47.93	50.53	57.80	7.27	24.07
	Im	70	57.28	30.22	50.09	49.99	60.69	68.44	7.76	24.03
Ee	05	18	54.22	28.39	51.00	50.44	62.00	69.11	7.11	20.19
Ee	06	26	52.78	31.15	50.77	52.24	58.50	66.42	7.92	24.62
Ee	07	26	51.25	29.54	50.15	50.12	60.15	67.69	7.54	20.17
Ee	08	22	58.80	31.50	53.00	50.62	60.86	68.73	7.86	22.28
	Ee	92	54.07	30.24	51.18	50.90	60.22	67.86	7.64	21.97
Es	09	18	56.60	27.17	44.69	48.94	52.67	61.61	8.94	22.39
Es	10	22	53.58	29.05	47.41	49.40	58.23	69.09	10.86	22.75
Es	11	17	58.00	30.18	52.41	50.76	59.82	68.47	8.65	18.55
Es	12	16	50.57	31.31	40.47	48.94	53.50	57.19	3.69	21.93
	Es	73	54.73	29.34	46.51	49.51	56.19	64.49	8.30	21.73
	sk	235	55.24	29.95	49.41	50.19	59.11	66.99	7.88	22.55

School Class Means in Certain Variables. ak; N = 12.

Method	School Class No.	N	DBA	Grades	Std Test	PACT	Pre-test	Post-test	Pro-gress	Pupil Attitude
Im	21	13	42.75	23.75	43.46	45.23	36.00	37.23	1.23	23.54
Im	22	12	39.70	20.25	31.09	43.36	29.67	33.42	3.75	19.91
Im	23	11	45.33	24.00	40.36	45.10	35.18	37.82	2.64	21.20
Im	24	14	43.92	25.29	37.86	41.00	31.79	34.57	2.79	21.43
	Im	50	42.91	23.39	38.39	43.65	33.12	35.70	2.58	21.60
Ee	25	8	45.38	20.63	38.25	43.25	32.13	36.38	4.25	17.29
Ee	26	14	36.32	21.43	31.23	37.14	28.93	29.71	0.79	18.92
Ee	27	18	40.91	22.50	37.00	42.07	28.72	31.22	2.50	19.33
Ee	28	9	46.22	25.00	41.11	40.89	36.67	39.67	3.00	25.29
	Ee	49	41.65	22.35	36.42	40.54	30.80	33.18	2.39	19.86
Es	29	19	48.18	25.89	44.82	45.44	32.32	39.74	7.42	20.53
Es	30	15	45.63	24.40	39.80	45.46	29.07	35.13	6.07	24.07
Es	31	12	43.92	24.27	42.67	45.58	31.42	36.25	4.83	24.08
Es	32	7	45.17	20.40	42.71	44.33	28.43	32.71	4.29	20.29
	Es	53	45.76	24.54	42.55	45.34	30.68	36.72	6.04	22.34
	ak	152	43.37	23.43	39.18	43.19	31.52	35.24	3.72	21.34

		Pre-test						Post-test							Progress						
		2	3	4	5	6	Tot	1	2	3	4	5	6	Tot	1	2	3	4	5	6	Tot
Pre-test	1	.354	.438	.448	.506	.540	.694	.515	.332	.381	.447	.511	.504	.600	-.640	-.034	-.186	.041	-.030	-.044	-.225
	2		.334	.414	.467	.534	.690	.267	.516	.182	.427	.392	.470	.539	-.150	-.517	-.247	.055	-.130	-.081	-.325
	3			.386	.472	.460	.641	.371	.378	.629	.431	.529	.460	.610	-.150	.033	-.658	.092	.039	.002	-.110
	4				.451	.512	.798	.397	.352	.282	.654	.467	.457	.644	-.137	-.076	-.215	-.339	-.011	-.069	-.340
	5					.713	.792	.401	.436	.336	.465	.715	.634	.682	-.198	-.047	-.273	.058	-.438	-.100	-.262
	6						.823	.451	.489	.308	.536	.654	.711	.729	-.190	-.063	-.285	.077	-.126	-.375	-.236
	Tot							.527	.550	.442	.686	.715	.708	.847	-.292	-.163	-.384	-.066	-.155	-.146	-.356
Post-test	1								.291	.415	.448	.499	.452	.650	.330	.015	-.068	.100	.098	.004	.170
	2									.327	.387	.421	.512	.664	-.105	.466	-.162	.076	-.051	.033	.152
	3										.314	.389	.391	.556	-.047	.138	.172	.065	.044	.112	.163
	4											.509	.516	.831	-.091	-.055	-.242	.490	.025	-.022	.194
	5												.671	.790	-.115	.016	-.293	.095	.316	.026	.070
	6													.800	-.150	.026	-.203	.116	.003	.385	.099
	Tot														-.079	.107	-.234	.291	.091	.098	.195
Progress	1															.051	.144	.045	.121	.052	.401
	2																.093	.019	.084	.117	.487
	3																	-.054	-.007	.106	.297
	4																		.043	.052	.633
	5																			.169	.446
	6																				.440
	Tot																				

sk, N = 235

		Pre-test						Post-test							Progress						
		2	3	4	5	6	Tot	1	2	3	4	5	6	Tot	1	2	3	4	5	6	Tot
Pre-test	1	.172	.186	.184	.194	.130	.495	.401	.160	.230	.313	.392	.300	.491	-.414	.010	.127	.125	.293	.227	.095
	2		-.055	.180	.121	.180	.533	.065	.434	-.048	.128	.138	.138	.245	-.075	-.525	-.011	-.031	.059	.028	-.251
	3			.194	.233	.143	.451	.393	-.036	.630	.178	.342	.366	.477	.240	.017	-.094	.001	.202	.286	.123
	4				.230	.175	.797	.324	.071	.280	.267	.145	.024	.363	.172	-.101	.185	-.557	-.025	-.087	-.379
	5					.141	.470	.334	.112	.205	.119	.547	.256	.376	.174	-.007	.053	-.078	-.208	.173	-.023
	6						.325	.194	.247	.179	.139	.220	.357	.330	.087	.065	.100	-.018	.138	-.269	.072
	Tot							.481	.265	.414	.353	.427	.299	.629	.077	-.248	.124	-.323	.101	.101	-.259
Post-test	1								.083	.427	.157	.370	.283	.582	.668	.018	.193	-.120	.149	.169	.222
	2									.003	.063	.182	.101	.413	-.047	.539	.037	-.002	.118	-.053	.238
	3										.192	.324	.287	.612	.238	.047	.714	-.055	.204	.182	.329
	4											.194	.204	.711	-.099	-.061	.086	.652	.126	.121	.517
	5												.342	.583	.050	.044	.107	.053	.705	.212	.280
	6													.454	.038	-.034	.038	.157	.183	.804	.253
	Tot														.180	.162	.354	.328	.363	.258	.587
Progress	1															.026	.089	-.220	-.089	-.016	.143
	2																.045	.027	.057	-.077	.460
	3																	-.071	.080	-.025	.311
	4																		.129	.173	.744
	5																			.101	.347
	6																				.215
	Tot																				

ak, N = 152